AD-778 312

FEASIBILITY OF AUTOMATED ADAPTIVE GCA (GROUND CONTROLLED APPROACH) CONTROLLER TRAINING SYSTEM

Robert L. Feuge, et al

Logicon, Incorporated

Prepared for:

Naval Training Equipment Center

April 1974

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

UNCLASSIFIED
Security Classification

AD 778312

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Logicon, Inc., P. O. Box 80158 1075 Camino Del Rio South, San Diego, Calif. 92138	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP NA

3. REPORT TITLE

FEASIBILITY OF AUTOMATED ADAPTIVE GCA CONTROLLER TRAINING SYSTEM

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final Report Dec 1972 – Jan 1974

5. AUTHOR(S) (First name, middle initial, last name)

Dr. Robert L. Feuge
Dr. John P. Charles
Mr. Ralph Miller

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1974	87	234

8a. CONTRACT OR GRANT NO. N61339-73-C-0079/ARPA Order No. 2310 b. PROJECT NO. c. Task No. 3753-01 d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report: NAVTRAEQUIPCEN 73-C-0079-1 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Naval Training Equipment Center Orlando, Florida 32813

13. ABSTRACT

An analysis of the conceptual feasibility of using automatic speech recognition and understanding technology in the design of an advanced training system was conducted. The analysis specifically explored application to Ground Controlled Approach (GCA) controller training. A systems engineering approach was followed to determine the feasibility of such a system. Design features were developed including training requirements and constraints. An evaluation of the state-of-the-art of speech understanding systems was conducted.

The results of the study indicate that the technology for automatic speech recognition and understanding is adequate to warrant design and construction of a feasibility demonstration model for the precision approach radar (PAR) phases of GCA controller training. As conceived, the system would accept student controller speech and convert it to functional 'commands' to drive the simulated radar return of an aircraft on a display in a manner much as the controller guides an actual aircraft on final approach to landing. Other design features would include objective performance measurement and adaptive syllabus control for increasing the difficulty of the training problem as a function of the student's performance. It is recommended that a technical feasibility demonstration be implemented.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Adaptive Training						
Automated Training						
Automated Speech Understanding						
GCA Controller Training						

UNCLASSIFIED

Security Classification

FOREWORD

This report describes the first steps in a program to develop techniques for improving quality control while reducing costs in the training of Navy personnel in certain of the skills needed to perform control jobs such as Ground Controlled Approach Controller, Ground Controlled Intercept Controller, and Radar Intercept Officer. These jobs have in common the need to use restricted, stylized speech to guide a recipient man-machine system to a well-defined goal. Application of the advanced technology of machine understanding of spoken commands in combination with the previously developed technologies of automated adaptive training should make it possible to realize important savings in manpower and other energy resources.

The outcome of the initial effort, as reported in this document, is a functional design for an experimental laboratory version of an advanced technology controller training system for GCA Controllers. This design is being implemented for use in the Human Factors Laboratory of the Naval Training Equipment Center. When installed the experimental system will be subjected to a period of test and evaluation to determine the technical feasibility of the concept.

Ira Goldstein

IRA GOLDSTEIN
Scientific Officer

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SECTION I

INTRODUCTION

In 1969, the Naval Training Equipment Center (NAVTRAEQUIPCEN) initiated a program to demonstrate that the effectiveness of training devices can be increased by the application of advanced technology. Of specific interest were the rapidly developing technology in computer sciences (especially programming) and in psychology (adaptive training and performance measurement). This study is directed to advances in automatic speech recognition (ASR) technology. ASR has developed from an interdisciplinary approach which includes computer sciences and psychology as well as other areas such as linguistics, phonetics, and artificial intelligence. Speech understanding encompasses the identification and understanding of human speech sounds by machine.

This is the final report of a project designed to establish the conceptual feasibility of exploiting these technologies for training communications skills. Adaptive training and objective performance measurement have been demonstrated earlier.[1] The major impetus of this report was the study of ASR technology and the analysis of the feasibility of a complete training system. The training requirement selected for demonstration was that of developing the skills and vocabulary needed by GCA controllers for final approach control.

The report includes a functional design for a precision approach radar (PAR) training system which uses ASR technology to recognize controller commands. After recognition (identification), the commands are converted to functional outputs reflecting the input of the student controller. The functional system includes other advanced training techniques such as automated-adaptive training.

Such a system should reduce the variability introduced by the many manual functions inherent in existing training equipment. The system also provides objective measures of performance and, by incorporating adaptive techniques, allows each student to progress through the program at his most effective rate.

---

[1] Charles, J. P., Johnson, R. M. and Swink, J. R. Automated Flight Training (AFT) GCI/CIC Air Attack. Technical Report NAVTRAEQUIPCEN 72-C-0108-1. July, 1973. Naval Training Equipment Center. Orlando, Florida.

SECTION II

STATEMENT OF THE PROBLEM

GENERAL

The successful combining of equipment and techniques into one functional system cannot be based solely on the success of individual or separate technical demonstrations. A period of conceptual analysis preceding technical implementation is required to determine the feasibility of combinations. A major problem is in the interfacing of system components. It is especially critical when the technologies represented are from widely separated and distinct disciplines. Therefore, the importance of isolating and analyzing conceptual feasibility increases with system complexity and particularly with rapidly advancing technology. Conceptual feasibility must be established, either analytically or empirically, before the demonstration of technical feasibility can be undertaken with reasonable confidence.

STATEMENT OF THE PROBLEM

The present study is primarily an analysis of conceptual feasibility of an advanced system for GCA controller training. It required exploring the possibility of combining automatic speech understanding with automated-adaptive training to form an advanced computer-based training system. A comprehensive review of speech understanding techniques consumed a major portion of the early analyses, since the literature was scattered throughout several academic and applied disciplines. Other portions of the study were dedicated to the application of these techniques and to the development of a functional design for the proposed system. Laboratory investigations of 1) speech timing patterns and 2) feasible improvements to current speech recognition devices were conducted.

Implementation of a demonstration model of the complete system and its test and evaluation remain to be accomplished.

SECTION III

METHOD

A systems engineering approach was taken to ensure an efficient and thorough evaluation of conceptual feasibility. Of particular importance were the analyses of PAR training requirements and the state-of-the-art necessary to establish the feasibility that advanced technology could be applied to training requirements. In-depth analysis of these factors was necessary to provide the study with sufficient scope to be representative of a typical training problem as well as predictive of feasibility and applicability of the application.

The project began with a constraint and training requirement analysis and concluded with a system functional design. Fourteen study tasks were identified as shown in Figure 1. The first five are oriented toward definition of the problem while Tasks 6 through 10 are oriented toward system requirements. The last four (Tasks 11 through 14) are system functional design tasks.

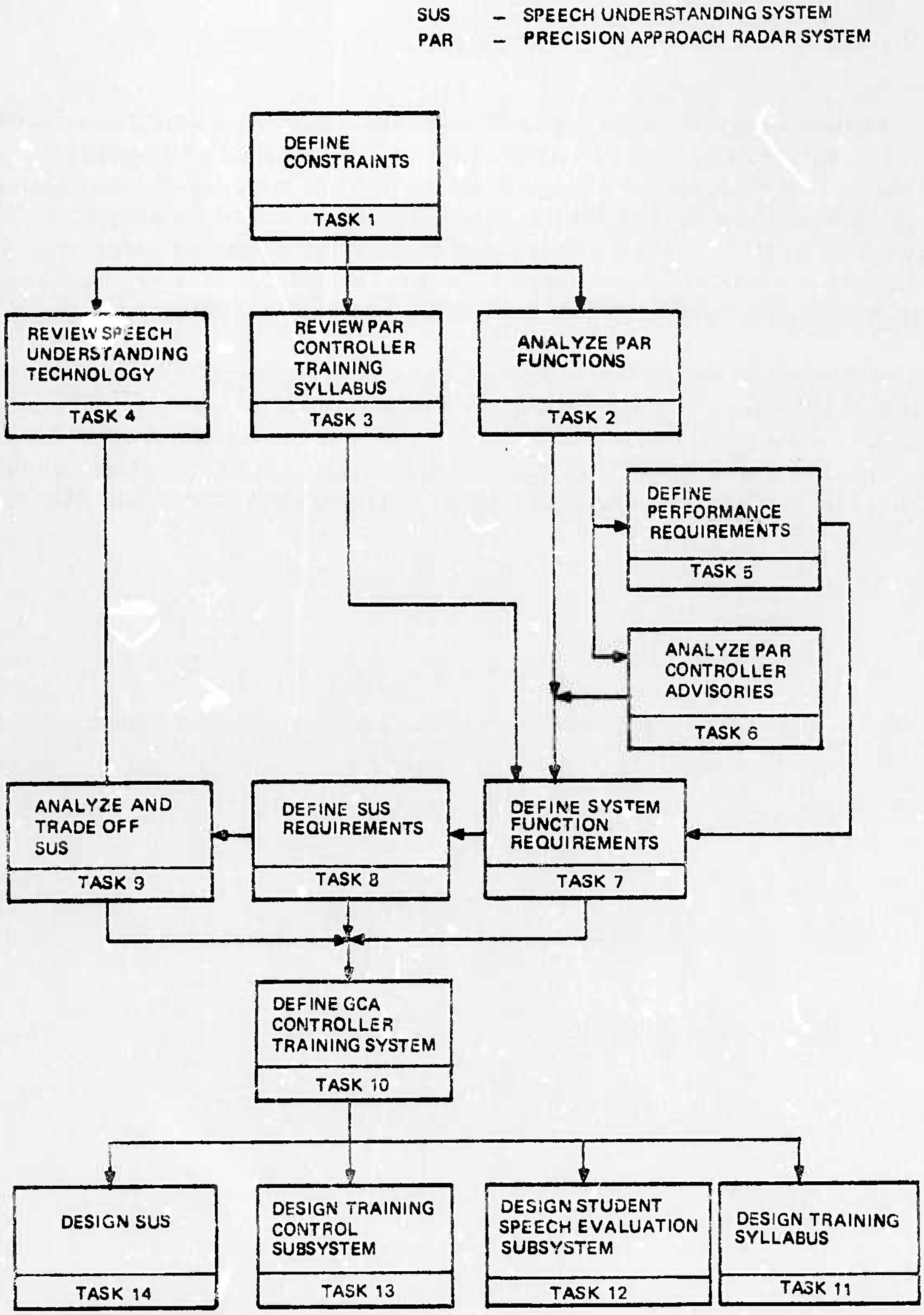


Figure 1. Flowchart of Study Tasks

SECTION IV

RESULTS

This section reviews the tasks outlined in figure 1. The results led to the final design specification. In all cases, the results refer to a 'demonstration system' design in contrast to an operational training system design. The requirements for the demonstration system are, however, representative of fully implemented systems and stress the risk areas that require actual demonstration of technical feasibility. The first tasks completed in the study were the detailed definition of the training problem including analyses of constraints, training requirements, and available technology.

TASK 1 – DEFINE CONSTRAINTS

The first task included defining constraints for the study and for the advanced training system. Constraints on the study were:

a. Time frame – The study was constrained to 1 year in which both a survey of existing speech recognition technology and a feasible system functional design were to be completed.

b. Current speech technology – Since the goal was an applied training system using advanced but available equipment, the study precluded design or development of a new recognition device. However, it did not rule out feasible improvement to existing devices.

Constraints on the training system originated from three sources:

a. Training syllabus characteristics.

b. Equipment characteristics and limitations.

c. Implementation limits, both in terms of supporting equipment and students.

These constraints were identified as follows:

a. Vocabulary and phraseology should be limited to the precision approach phase of GCA controller training, as defined by the Syllabus for the Navy GCA Controller Training Course, FAA Instruction 7110.8C, and standard Navy approach procedures.

b. Simulated approaches (student problems) should be standard PAR approaches (surveillance and no-gyro approaches need not be provided in the demonstration since they are, in effect, a simple subset of the precision approach).

c. Existing speech understanding technology requires 'training' of the system; i.e., each person using the system speaks each phrase into the system 5 to 10 times. The system stores individual speaker characteristics and adjusts the recognition threshold. The student must try to use the same speech patterns (e.g., intonation, inflection, stresses, and pauses) during training that he will use in system operation or test.

d. The demonstration system should be compatible with conventional computer systems such as the PDP-9 computer system at NAVTRAEQUIPCEN to be implementable.

TASK 2 – ANALYZE PAR FUNCTIONS

On each final approach, the PAR controller must perform the following functions:

a. Provide published decision height, if requested.

b. Issue glidepath intercept notification.

c. Issue instruction to begin descent.

d. Issue glidepath advisories.

e. Issue course advisories.

f. Issue range advisories.

g. Issue decision height advisory.

h. Issue clearance information.

i. Issue surface wind information.

j. Issue position advisories.

k. Issue instructions to execute missed approach.

l. Counter system malfunctions by corrective action.

m. Issue 'handoff' instructions.

To discharge these 13 functions properly, the PAR controller can use approximately 80 basic phrases specified in FAA Instruction 7110.8C. These are listed in appendix A along with the criteria for issuing each advisory. A time-line analysis for a typical PAR approach was conducted from recorded approaches made at Naval Air Station Miramar and is presented in appendix B. It can be seen that transmissions on final approach occur with less than a 5-second pause, and have a definite rhythm.

TASK 3 – REVIEW PAR CONTROLLER TRAINING SYLLABUS

The GCA controller training syllabus was reviewed to ensure that development of the demonstration system would be compatible with actual operational training requirements. As part of the review, several Navy facilities were visited, including the GCA school at the Naval Air Technical Training Center (NATTC), Glynco, Georgia and the GCA unit at the Naval Air Station Miramar, California. Navy personnel were interviewed regarding any details of training not stressed in the syllabus but which were operationally important. GCA controllers were also observed conducting approaches, and audio tape recordings were made of final approaches. These observations and recordings provided important additional information. The controllers were particularly helpful in providing message timing and priority information.

The data were used to identify training functions. The basic training requirements identified included developing the following vocabulary and procedural skills.

a. Use of GCA vocabulary.

b. Use of standard and approved PAR procedures.

c. Control of aircraft under various weather conditions.

d. Control of aircraft with different flight characteristics.

e. Control of aircraft with various pilot characteristics.

f. Control of aircraft with system malfunctions or emergencies.

g. Control of aircraft under various traffic loads.

TASK 4 – REVIEW SPEECH UNDERSTANDING TECHNOLOGY

To determine the 'state-of-the-art' of automatic speech technology, relevant literature in the fields of computer science, engineering, and linguistics, from early 1950 to the present, were reviewed. Major emphasis was given

to the period from 1965 to the present since most of the technical advances have occurred in this time frame.

The library search uncovered over 150 references to automatic speech recognition, analysis, synthesis, and understanding. A further search by the National Technical Information Service (NTIS) yielded 100 references, although some duplicated the 'academic' literature. The extensive bibliography developed, which contains over 200 entries, is included in this report.

A concurrent review of industry capabilities revealed that at least four companies presently build or are actively planning to build speech recognition devices or equipment. A review of the devices is presented in appendix C.

TASK 5 – DEFINE PERFORMANCE REQUIREMENTS

GENERAL. Technical manuals point out that radar provides the most precise means of obtaining information to guide an aircraft through an instrument approach to landing. The radar systems most commonly used are called the Ground Controlled Approach (GCA) system. It is composed of two subsystems: surveillance radar and a precision approach radar (PAR). Both subsystems provide azimuth and range information. The PAR system also provides height information on the landing aircraft. The controller must determine the relative position of the aircraft and advise the pilot through the approach. The advisories are a set of short, well-defined phrases designed to minimize ambiguity and misinterpretation.

Of the two subsystems, PAR presents the greater challenge to controller training. This stems from the fact that the final approach controller must determine and transmit glide slope, as well as heading and range information, with no pause greater than 5 seconds. If the pause is longer, the pilot assumes a loss of communications and executes a 'missed approach.' Landing clearance, wind, and position advisories must also be given at specific points in the approach. The major factors contributing to task difficulty include:

a. Decreasing safety tolerances with range.

b. Vocabulary constraints.

c. Weather.

d. Aircraft performance and type.

PRESENT PAR TRAINING. GCA controller instruction is currently conducted at the Naval Air Technical Training Center at Glynco, Georgia. The GCA training program utilizes training devices such as illustrated on figure 2. The student is provided a simulated GCA control console which includes communication equipment. For PAR controller instruction, the display presents azimuth, elevation, and range. The student PAR controller transmits advisories to an 'acting pilot' who 'flies' the simulated aircraft. Aircraft position changes which occur as a result of the 'acting pilot's' flight response are displayed on the student's radar display through a video simulator. The instructor supervises training sessions, subjectively evaluates student performance, and implements the overall training plan by altering PAR conditions to present a variety of problems to the student. Two men are normally required to teach one student controller.

The typical existing training system also lacks the following basic training capabilities:

a. Objective performance measurement.

b. Realistic aircraft performance.

c. Individualized instruction.

d. Extrinsic real-time feedback for the student.

SPECIFIC PAR REQUIREMENTS. As already discussed, the PAR subsystem includes radar displays of range-azimuth and range-elevation. These displays are used by the controller to guide the aircraft on the landing approach from approximately the final approach fix to the point where the aircraft is over the landing threshold. The PAR controller uses specified phrases to advise the acting pilot of his position with respect to the glidepath and the extended runway centerline. Each phrase has a unique criterion for its use (refer to appendix A).

PAR training requirements involve both content and timing of messages. Content is a matter of using correct vocabulary and phraseology in message transmission. Timing refers to when a message is to be sent and includes both message spacing and message priority. The training problem is to teach the student what he should say in specific cases, and when he should say it. For example, when an aircraft radar return is about two-thirds above the glidepath cursor, the correct advisory is:

'slightly above glidepath,' and is given as message priority permits.

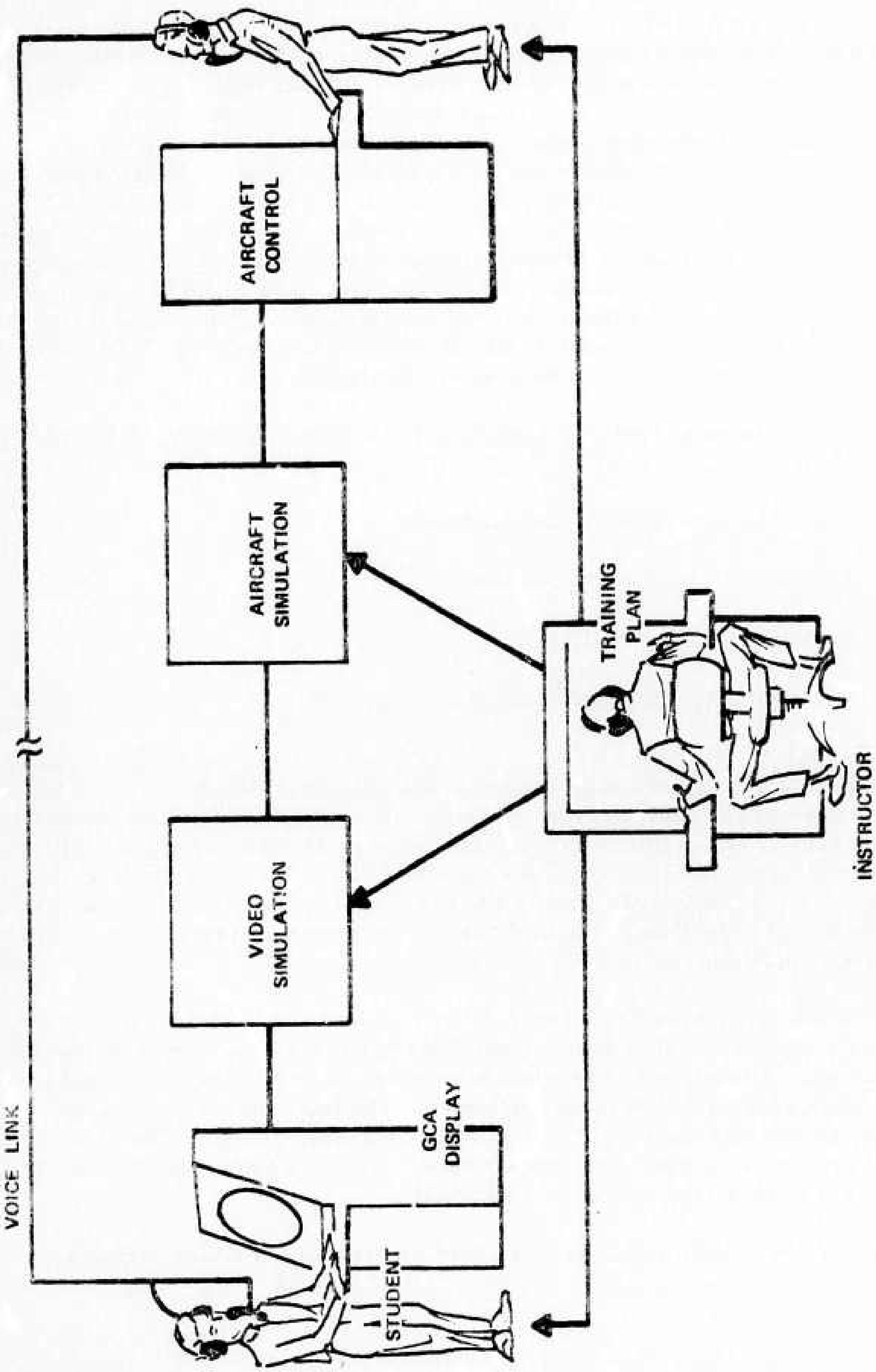


Figure 2. Typical GCA Controller Training System

If the aircraft return is completely above and separated from the glidepath cursor, the correct advisory is:

'well above glidepath.'

Similar criteria are used for heading advisories.

Under normal circumstances, the controller should achieve a 75 percent glidepath and 25 percent course advisory ratio. Range information (in miles to touchdown) should be given once per mile, and surface wind information and clearance to land are given at about 3-miles.

Based on the Navy GCA Controller Course Syllabus (Lesson Guide 2.2.4.1), the controller is evaluated in terms of:

a. Control accuracy.

b. Clarity of instruction.

c. Conformance to standard phraseology.

d. Conformance to standard procedure.

According to the syllabus, "Any tendency of controllers to give erroneous advisories or instructions or to become confused is considered a factor (in the evaluation of competence) regardless how well the equipment works. Simulated emergencies – will be used to determine controller competence in emergencies."

Thus, both the format and content of advisories as well as procedures used by the students are important and must be considered.

TASK 6 – ANALYZE PAR CONTROLLER ADVISORIES

The review of the existing GCA Syllabus at NATTC, Glynco, Navy, and relevant FAA instructions revealed approximately 80 basic phrases for PAR controlling (refer to appendix A). A detailed analysis of the advisories was conducted. It revealed several regularities which are significant for automatic recognition. They are:

a. The advisories can be divided into three categories: glidepath, heading (course), and ancillary information.

b. Advisories in each of the three categories begin with unique sets of words which are nearly mutually exclusive (refer to table 1).

c. The syntax of phrase composition is highly regular.

d. There is considerable redundance of information within phrases.

TABLE 1. CATEGORY DISTINCTIONS BY FIRST WORD

Glidepath	Heading	Information
Approaching	Turn	Cleared
Begin	Heading	Wind
On	On	At
Going	Going	No
Slightly	Slightly	(1 – 6)
Well	Well	Over
Coming	Right	Contact
Above	Left	
Below		
Execute		

Note: Only 'on,' 'going,' 'well,' and 'slightly' are common to different categories. By analyzing the second word (following 'well' and 'slightly'), the heading and glidepath phrases can be distinguished. 'Going' requires at least two subsequent words before the differentiation can be readily made.

PHONETIC ANALYSIS. Most approaches to automatic speech recognition attempt to analyze speech into phonemic or subphonemic segments. Therefore, a phonetic analysis was completed on basic or key PAR words that distinguish phrases from one another. As shown in table 2, there is little similarity between phonetic representations for key words. Of those words having similarity, none were classified as minimal pairs. A minimal pair refers to two words that differ by only one phoneme. Only the words 'heading' and 'holding' contain as many as four identical phonemes. The problem was minimized when systems test indicated that heading and holding could probably be readily differentiated on other criteria. (Most automatic recognition techniques concentrate on vowels, which is precisely where the two words heading and holding differ most; i. e., e vs. ow.)

TABLE 2. PHONETIC COMPARISONS OF PAR WORDS

Words	Phonemes by Position									
	1	2	3	4	5	6	7	8	9	10
Above	a	b	ə	v						
Approaching	a	p	r	o	w	č	i	ŋ		
Assigned	a	s	s	a	y	n	d			
At	æ	t								
Begin	b	i	y	g	i	n				
Below	b	ɨ	y	l	o	w				
Cleared	k	l	ɨ	y	r	d				
Course	k	o	w	r	s					
Decision	d	i	y	s	i	š	ə	n		
Eight	e	y	t							
Four	f	o	w	r						
Five	f	a	y	v						
Glidepath	g	l	a	y	d	p	æ	θ		
Going	g	o	w	i	ŋ					
Heading	h	e	d	i	ŋ					
Holding	h	o	w	l	d	i	ŋ			

TABLE 2. PHONETIC COMPARISONS OF PAR WORDS (Cont)

Words	Phonemes by Position									
	1	2	3	4	5	6	7	8	9	10
Left	l	e	f	t						
Nine	n	i	y	n						
One	w	ə	n							
One-half	w	ə	n	h	a	f				
Over	o	v	ə	r						
Right	r	a	y	t						
Runway	r	ə	n	w	e	y				
Seven	s	e	v	ə	n					
Six	s	i	k	s						
Slightly	s	l	a	y	t	l	ɨ	y		
Surface	s	ə	r	f	æ	s				
Take	t	e	y	k						
Three	θ	r	i	y						
Two	t	u	w							
Well	w	e	l							
Zero	z	ɨ	y	r	o					

PAUSE ANALYSIS. A major problem encountered in automatic recognition of speech is the detection of the beginning and ending of an utterance. The acoustic analog of language is a more or less continuous function. Yet, it represents discrete units of meaning (words, morphemes). Thus an automatic speech recognizer must be designed to break (segment) the acoustic analog into these discrete units at the proper places.

The approach typically taken is to detect pauses of predetermined lengths (e. g., 20 milliseconds). The pause is defined as a period of time in which 'significant' acoustic energy is not present. The detection of pauses is complicated by the fact that some intraword pauses have durations that approximate interword or interphrase pause lengths. For example, 'stops'

such as b, p, t, and g all have pauses created by the constriction of the oral cavity. One could mistakenly segment the word "above" into two words, 'a' and 'bove' because of the bilabial stop associated with the 'b.' Should this happen in PAR controlling, the phrase 'slightly above glidepath' would be segmented into 'slightly a' and 'bove glidepath,' neither of which conveys the full meaning.

Since estimates of the durations of intraword, interword, and interphrase pauses for GCA controllers were obviously required, audio tapes were made of PAR controllers during actual approaches. Samples of these tapes were then input into a Model 6061B Spectrum Analyzer and sonagrams made of the different phrases. (A sonagram is a visual frequency-time-amplitude display of the acoustical signal.) Figure 3 is a sample sonagram for the advisory 'two miles from touchdown.' Since the sonagram is a precise representation of sound over time, accurate calculation of pause durations was possible. For the samples used, intraword pauses averaged 0.034 second while interword pauses and interphrase pauses averaged 0.044 and 0.122 second, respectively. Distribution analysis showed some overlap.

These pause-length data were used to discuss the PAR phrase-understanding problem with manufacturers of speech recognition devices. Most existing devices require some sort of 'training'; i.e., analogs of the utterance of each speaker are stored for later comparison in the recognition process. The detection of pauses is not considered to be as difficult as it would be if the devices were designed for continuous free speech. Furthermore, the interphrase pause lengths are large (relative to intraphrase pauses). Therefore, discrimination of pauses should probably present little or no problem to the proposed system.

SPEECH UNDERSTANDING ACCURACY. Speech understanding subsystem (SUS) output will be utilized for three different functions. While 100 percent correct understanding of any student's advisory output is desirable, lesser accuracy can be tolerated, especially if a 'bootstrap' approach utilizing redundant information in the system is mechanized to support the particular requirement for each function. The three functions are: control of the pilot/aircraft simulation, evaluation of student's vocabulary, and evaluation of student's message. One of the most stringent requirements is imposed by heading advisories which utilize numerals for heading control. However, heading advisories should be confined to 3- to 5-degree heading changes. Thus ambiguities or decision confusions can be resolved in favor of the most logical choice. This same approach can be employed for other functions. The preliminary analyses have indicated redundant information will be available to aid in resolving recognition confusions or to resolve inconsistencies between speech understanding output and other system data so as to be meaningful for training purposes. Thus the approximately 90-percent accuracy achievable with state-of-the-art speech recognition equipment should prove adequate for demonstration purposes. Accuracy should approach 99 percent.

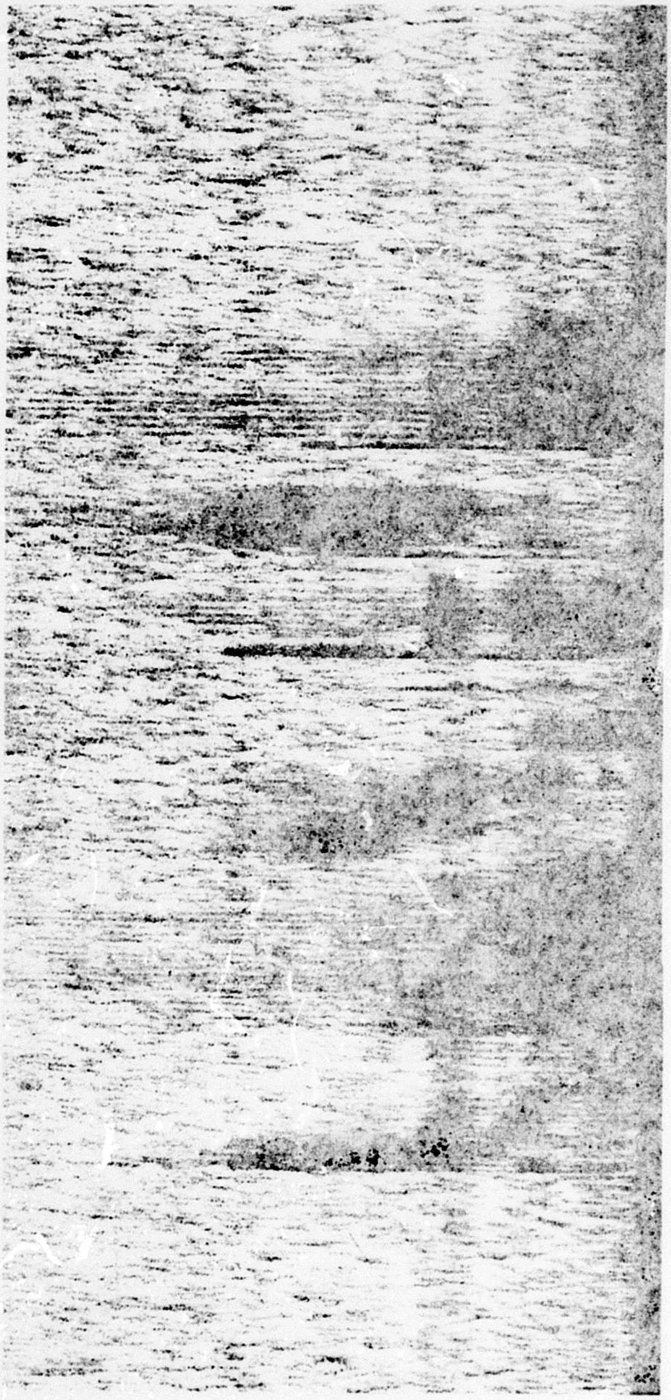


Figure 3. Sonagram of the Advisory 'Two Miles from Touchdown'

TASK 7 – DEFINE SYSTEM FUNCTION REQUIREMENTS

A function analysis was conducted and reviewed against system performance requirements. From this analysis, the following function requirements were derived:

a. Automatically 'understand' approximately 80 PAR phrases (each phrase composed of not more than nine words).

b. Convert the 'understood' phrase to a form usable for aircraft control.

c. Simulate 'pilot dynamics,' various aircraft types, and environment factors to create realistic training problems.

d. Simulate the PAR radar display.

e. Objectively measure and evaluate student performance.

f. Provide on-line performance feedback to the student.

g. Provide automatic-adaptive on-line syllabus control.

h. Provide a hardcopy printout of each student's performance.

i. Summarize and store training data for each student.

Figure 4 shows the functional flow of the system for one student.

The PAR functions were then analyzed in depth to isolate the training task and support functions required. As mentioned before, the Navy GCA training syllabus was reviewed and interviews with PAR controllers were conducted to establish the training requirement.

The major variables affecting PAR controllers were identified as:

a. Wind factors.

b. Type of aircraft.

c. Pilot variability.

d. Response lag.

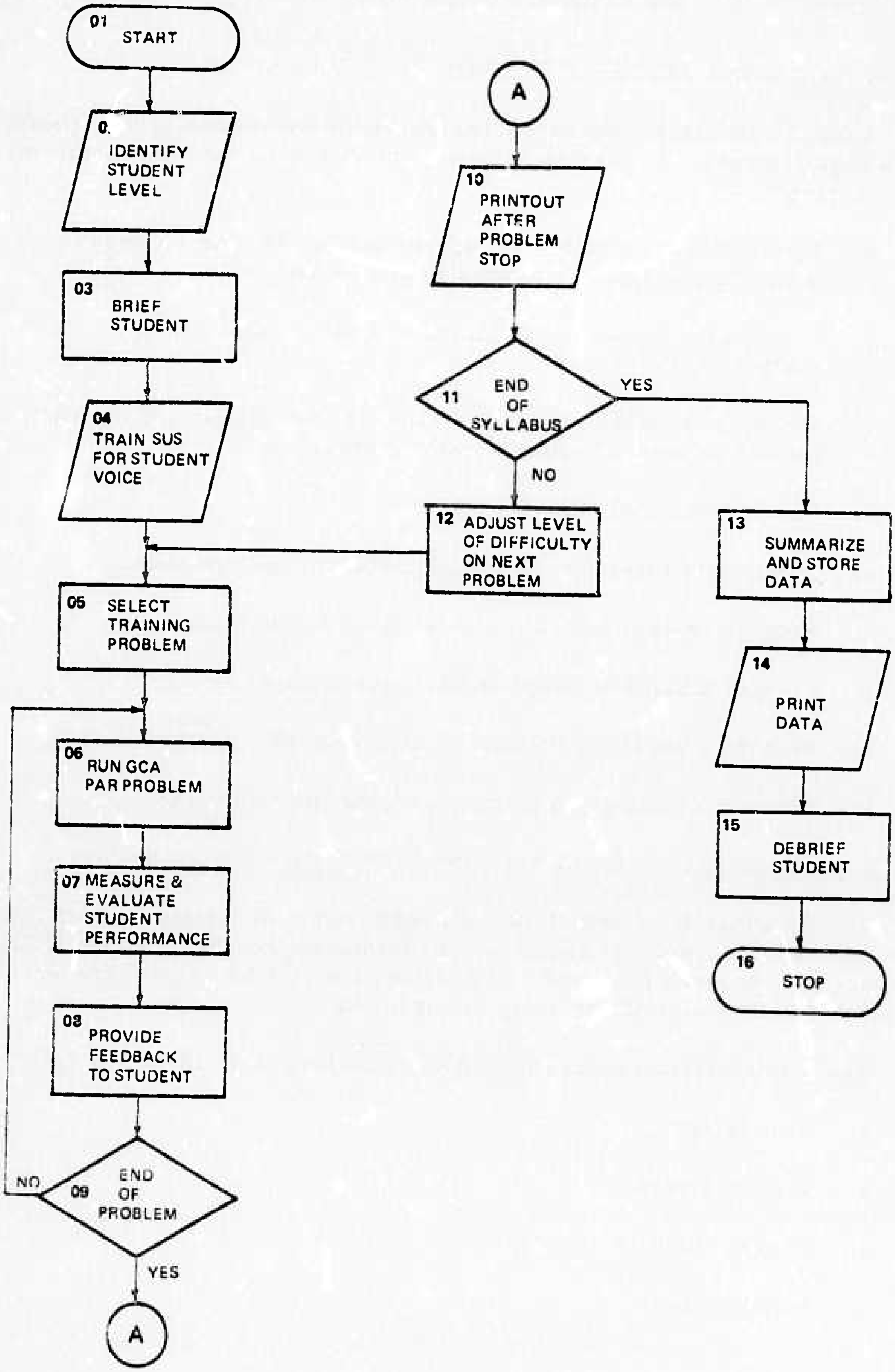


Figure 4. Training System Function Flow

e. Task loading.

f. Type of approach.

The information requirements of the controller were identified as:

a. Glidepath cursor.

b. Azimuth cursor.

c. Aircraft return (targets or pips).

d. Mile markers.

e. Safety limits.

f. Tower clearance.

g. Published decision height.

h. Runway in use.

i. Wind conditions.

Figures 5 and 6 depict the glide slope and approach course geometry for the PAR system.

A detailed analysis of the training steps involved was conducted to ensure exploration of all contingencies.

Figure 7 depicts the basic training session in a first-level flow diagram, and identifies the major support systems functions required.

TASKS 8 AND 9 – DEFINE SPEECH UNDERSTANDING SYSTEM REQUIREMENTS AND TRADE OFF SPEECH UNDERSTANDING SYSTEMS

A comprehensive review of the goals of machine speech understanding development was sponsored by the Advanced Research Projects Agency in the spring of 1970. The review was reported by Newell, et. al. (1971) and concluded that the following specification characteristics were a reasonable goal.

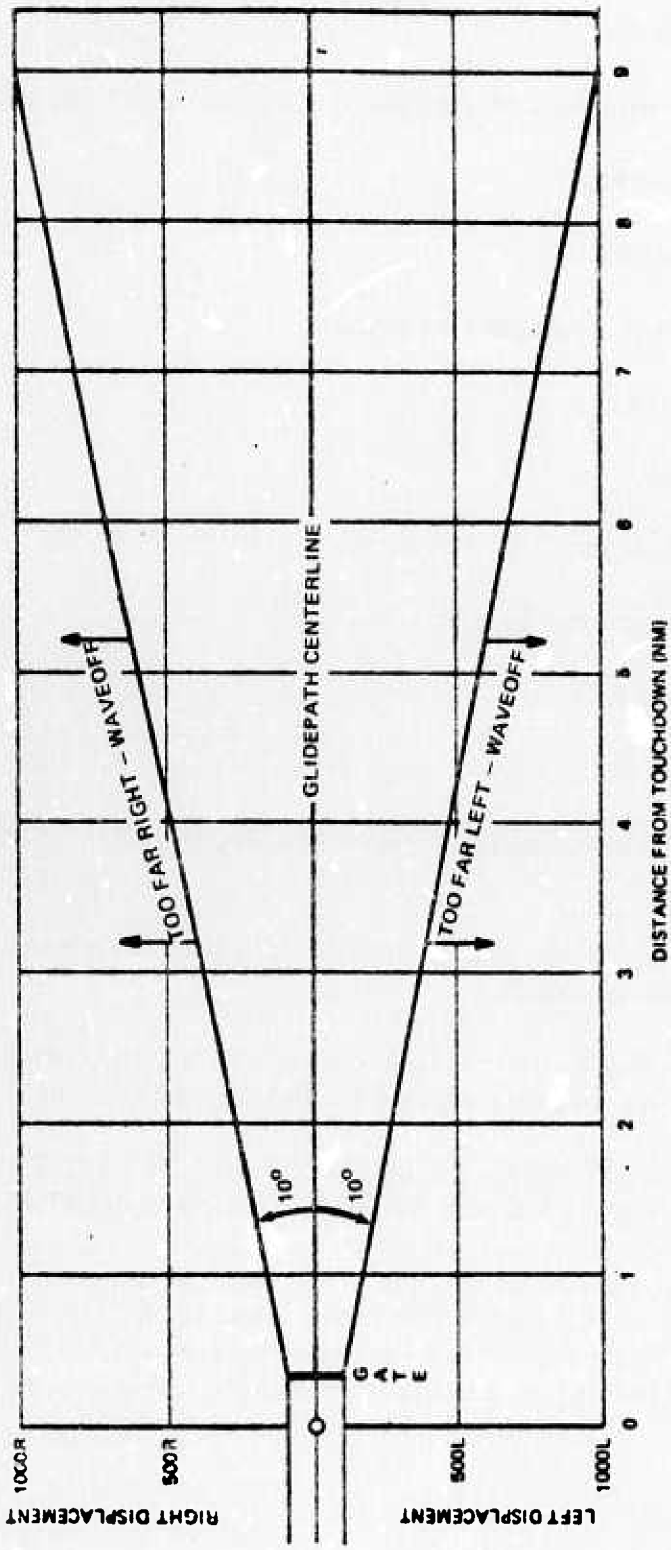


Figure 5. PAR Approach Course Geometry

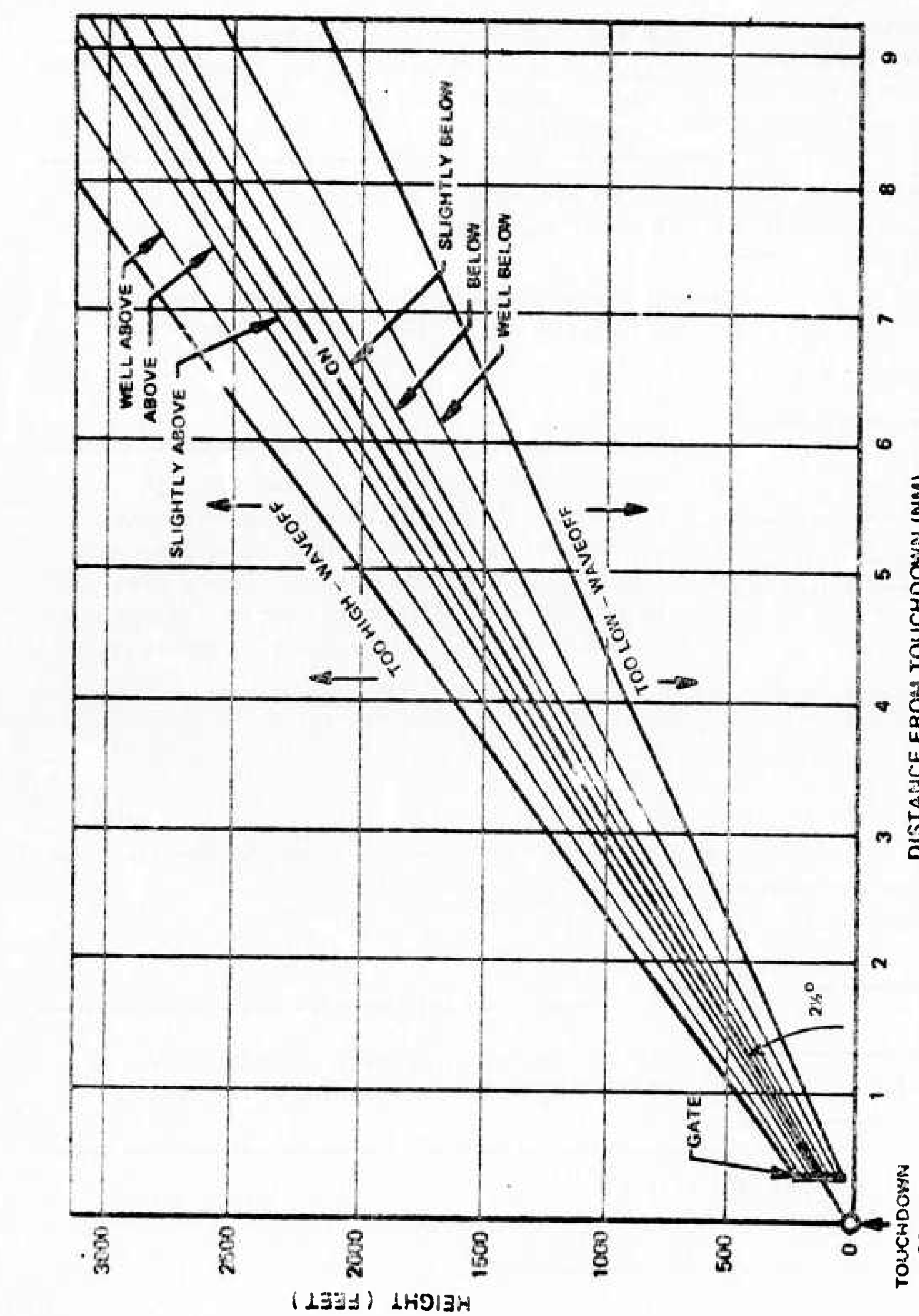
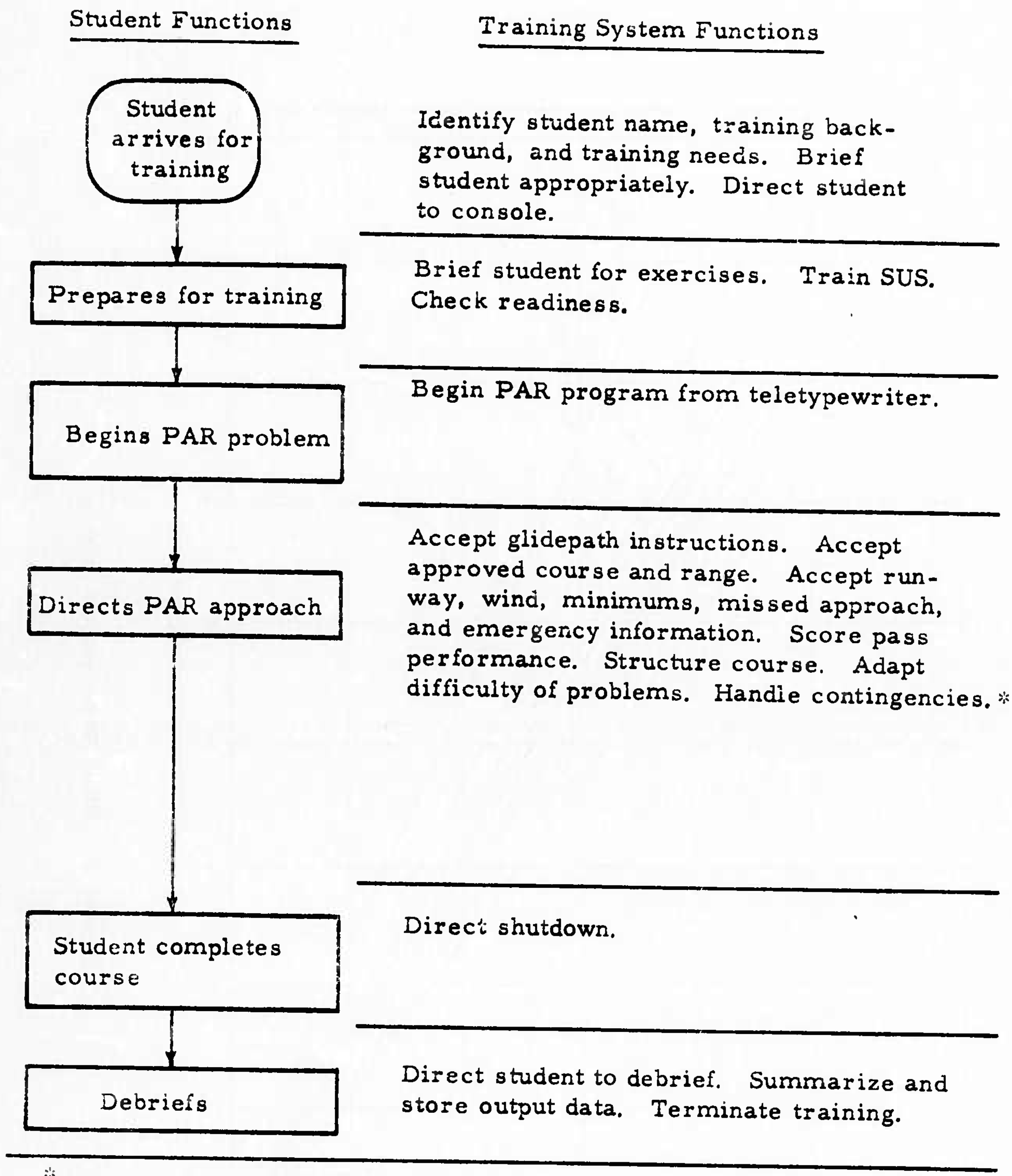


Figure 6. PAR Glidepath Geometry



*Major contingencies to consider:

a. Crash.
b. Student fails to take directed action.

Figure 7. PAR Training Functions

The system should:

(1) Accept continuous speech
(2) from many
(3) cooperative speakers of the general American dialect
(4) in a quiet room
(5) over a good quality microphone
(6) allowing slight tuning of the system per speaker
(7) but requiring only natural adaptation by the user
(8) permitting a slightly selected vocabulary of 1000 words
(9) with a highly artificial syntax
(10) and a task like the data management or computer status tasks (but not the computer consultant task)
(11) with a simple psychological model of the user
(12) providing graceful interaction
(13) tolerating less than 10 percent semantic error
(14) in a few times real time
(15) and be demonstratable in 1976 with a moderate chance of success.

These characteristics are accepted as a realistic goal and provide a standard against which the training requirements can be evaluated.

After reviewing the Newell characteristics, LOGICON conducted an analysis of PAR advisory phrases and compiled a similar list of characteristics for the speech understanding component for the GCA controller training system. The speech understanding component of the GCA controller training system should:

(1) Accept short phrases (noncontinuous speech)
(2) from many
(3) selected speakers of the general American dialect
(4) in a relatively noisy room
(5) over a good quality microphone
(6) allowing training of the system per speaker
(7) but requiring only natural adaptation by the user
(8) permitting a highly selected vocabulary of approximately 70 words
(9) with an extremely invariant and orderly syntax
(10) for a GCA task
(11) with a simple functional model of the pilot or aircraft
(12) providing graceful interaction
(13) tolerating less than 5 percent system response error
(14) in near real time
(15) and be ready for field testing in 1975.

Three of these specifications appear more stringent than their counterparts in the Newell, et. al. listing. The first of these, (3) above, refers to basic speech training and would involve control of individual voice characteristics such as dialect, timing, and patterns of inflections. Although limited, such control occurs initially by student selection. Controller training itself develops a certain discipline in speech habits. GCA phraseology is, of necessity, highly restricted and invariant when compared to free, connected speech. For example, vocabulary requirements for the GCA system are very limited (66 words).

The second, (4) above, refers to the ambient level of noise. The present system must deal with a certain amount of interference from a variety of related equipment including radios and other controllers. While noise in the operational controller environment itself cannot be appreciably reduced, it is subject to some control in the training environment. Fortunately, the speech analog itself is robust and provides a certain resistance to distortion. Major problems would be expected only if two or more voices were simultaneously picked up by the microphone, or if ambient noise was of sufficient volume to drown out the student's voice. These contingencies can be controlled with little effort by use of sensitive and directional microphones.

The third, (13) above, refers to tolerances for semantic error; that is, errors at the level of meaning. An allowable error rate of 10 percent is too high for the voice system, since the overall system response error rate should be 5 percent or less. Error tolerance was discussed briefly under Task 6. The Newell, et. al. system is based on continuous speech and a 1000-word vocabulary. Continuous speech in particular, and to a lesser degree the large vocabulary, precludes the possibility of a secondary scanning of the material within the specified real-time frame. The GCA system, concerned with short phrases, a highly selected, small vocabulary within a less stringent time frame, will allow for a secondary scan of the speech signal to be recognized. Newell was also concerned specifically with errors of meaning while the training system is concerned only with errors of response (understanding) to a specified input. Therefore, system accuracy greater than 95 percent is considered a reasonable and achievable goal.

In summary, none of the three major specification differences appears to be a limiting factor in training system design, nor do any of the other minor changes. Requirements for the speech understanding component of the training system are much less ambitious than the Newell requirements. Therefore, based on vocabulary and phrase requirements, the speech understanding component of the GCA controller training system appeared to be within the current state of the art. To verify this, LOGICON made a

preliminary survey of companies building speech recognition devices. The survey particularly emphasized the following objectives:

a. Recognition of phrases up to 4 seconds in length.

b. Ninety-five percent phrase accuracy.

c. Ability to handle multiple speakers and dialects.

d. Minimal system training or tuning for each different speaker.

The survey revealed four devices under development, three of which were operational at the time of the survey. One of the three was selected as the candidate for the GCA controller training system. Details of the survey and consequent trade-off are found in appendix C.

TASK 10 – DEFINE GCA CONTROLLER TRAINING SYSTEM

GENERAL. The initial system definition requires identification of high risk or critical design features. For the GCA controller training system, the critical features are those requiring a speech understanding or recognition capability. The design requirement can be limited to the PAR phase of operations since it is the limiting condition. Simulation of displays and training exercises need only be functionally realistic to establish technical feasibility.

DEFINITION. As presently conceived, the GCA controller training system design features four major subsystems. They are:

a. Adaptive syllabus control subsystem.

b. Student speech evaluation subsystem.

c. Training control subsystem.

d. Speech understanding subsystem.

The following four task descriptions (Tasks 11 – 14) include a discussion of each of these subsystems.

TASK 11 – DESIGN TRAINING SYLLABUS

A preliminary syllabus is based on the variables identified in Task 7. Of the six variables listed, four have been tentatively selected for use in the demonstration system. These four represent factors significantly affecting task difficulty (for the controller) during the PAR approach. The other two variables, task loading and type of approach, concern subject variables and

procedures not within the scope of the demonstration model (they primarily increase difficulty for the pilot and actually simplify the problem for the controller). These variables need not be included in the feasibility demonstration.

ADAPTIVE VARIABLES. For the demonstration, three values representing points along a continuum of difficulty were identified for each of the four variables previously described. Each point represents a level of difficulty. Further refinement of the values will occur during the early part of the demonstration as a function of the flight equations mechanized. Representative levels of difficulty for the demonstration are listed in the following:

a. Wind factors:

   1. Fifteen-knot head wind.
   2. Fifteen-knot cross wind.
   3. Thirty-knot (variable) cross wind, terrain turbulence.

b. Aircraft type. Three typical aircraft characteristics reflecting FAA categories were identified. These are, based on weight and speed:

   1. Less than 91 knots and 30, 000 pounds (category A).
   2. One-hundred twenty-one to 141 knots, 60, 000 to 150, 000 pounds (category B).
   3. One-hundred sixty-six knots, any weight (category E).

c. Pilot response lag (must be established after flight equations are identified):

   1. Average.
   2. Long.
   3. Short.

d. Pilot variability (must be established after flight equations are identified):

   1. 0. 5σ (standard deviation).
   2. 1. 0σ.
   3. 2. 0σ.

These 'pilot variables' reflect aircraft control characteristics as well as pilot control input. They must be quantified and checked during the early part of the demonstration phase.

SYLLABUS. Table 3 presents a feasible training syllabus containing 20 exercises. Each exercise was constructed by including one level of each of the four adaptive variables in an order which increased level of difficulty. The syllabus also reflects information obtained from GCA controllers at Miramar NAS. These controllers indicated that wind and aircraft factors are more salient than pilot factors, although all three contribute to difficulty of the approach from the controller's standpoint.

TABLE 3. PRELIMINARY TRAINING SYLLABUS

Sequence Number	Wind Factors	Pilot Response Lag	Pilot Variability	Aircraft Type
1	1	1	1	1
2	1	1	2	1
3	1	1	3	1
4	1	2	2	1
5	1	2	3	1
6	2	1	2	2
7	2	1	3	2
8	2	2	3	2
9	2	2	2	2
10	2	3	3	2
11	2	3	2	2
12	3	1	3	2
13	3	2	1	2
14	3	2	1	3
15	3	3	2	3
16	3	3	2	3
17	3	3	3	3
18	3	3	3	3
19	3	3	3	3
20	3	3	3	3

ADAPTIVE SYLLABUS CONTROL SUBSYSTEM (ASCS). The ASCS is presently conceptualized as consisting of the syllabus and the logic used to select the next training problem. The syllabus (table 3) is arranged in order of increasing difficulty. The selection of each succeeding problem will be based on the score attained on the previous problem. The student will be incremented more rapidly for good performance than for average performance. Poor performance will result in decrementing the level by one or more steps. The student moves through the program in accordance with the logic shown in figure 8 and graduates from it upon satisfactory completion of the most difficult problem. The logic is similar to that used in a previous program (ATE)[2]. It permits the student to complete the course in accordance with his ability.

TASK 12 – DESIGN STUDENT SPEECH EVALUATION SUBSYSTEM

The problem definition phase devoted considerable effort to identifying objective measures of student performance that might be used to reflect learning. Although present PAR controller performance measures are largely subjective, the syllabus used by the Navy provided guidelines for developing objective measures. Additional background data for performance measurement was collected at NAS Miramar. The end result was a list of nine measures which reflect the performance requirements discussed earlier. The measures include:

a. Percent correct advisories.

b. Ratio of glidepath to heading advisories.

c. Number of errors in phraseology.

d. Error of aircraft about glidepath.

e. Error of aircraft about centerline.

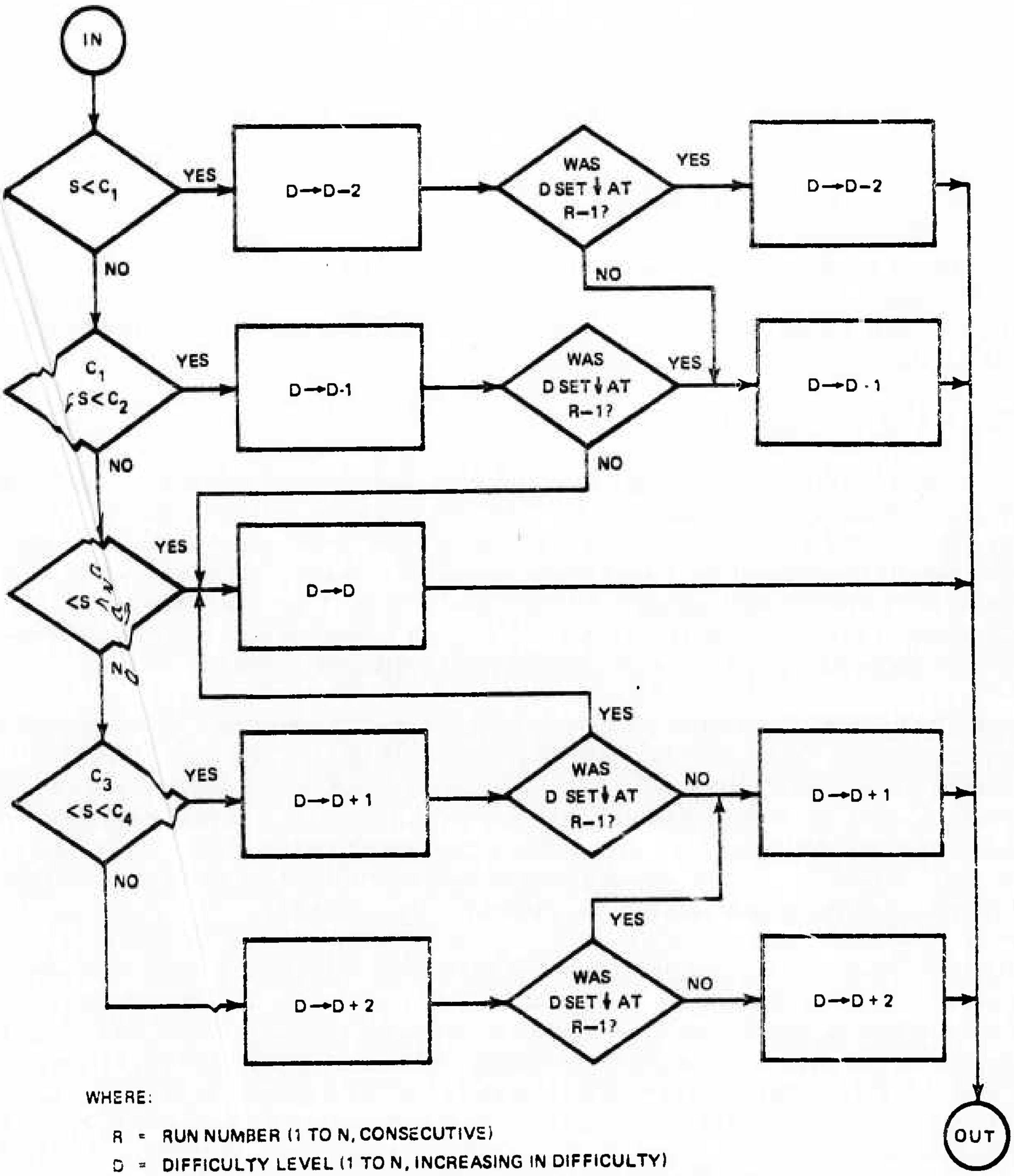
f. Number of advisories in each category (well above, below, right, left, etc.)

g. Number of 1-degree heading changes.

h. Time delay between advisories.

---

[2] Charles, J. P. and Johnson, R. M. Automated Training Evaluation (ATE). Technical Report: NAVTRADEVCEN 70-C-0132-1. January 1972. Naval Training Equipment Center. Orlando, Florida.



WHERE:

R = RUN NUMBER (1 TO N, CONSECUTIVE)

D = DIFFICULTY LEVEL (1 TO N, INCREASING IN DIFFICULTY)

S = SCORE FROM PERFORMANCE EVALUATION ROUTINE FOR PAST PROBLEM

C_N = CRITERIA FOR PERFORMANCE

Figure 8. Adaptive Logic Flowchart

i. Procedural errors or omissions; e.g., failure to issue:

1. "Begin descent."
2. Reference to touchdown.
3. Heading changes in 5-degree increments beyond 3 miles.
4. Two to 5-degree heading changes within 3 miles of touchdown.
5. Landing clearance.

Some or all of the nine measures will be combined to produce a composite score of PAR controller performance for use in the adaptive logic to adjust problem difficulty on successive runs. This score is referred to as 'C' in figure 8.

TASK 13 – DESIGN TRAINING CONTROL SUBSYSTEM

The training control subsystem is the vehicle for automated instruction. As such, it takes the selected problem from the adaptive syllabus control subsystem, initializes it, generates and controls the PAR display, converts the SUS output to aircraft parameters, implements adaptive variables, and, in general, controls the progress of the task. Figure 9 portrays the block diagram of these functions. Of these functions, two require further elaboration. These are aircraft simulation and PAR display simulation.

AIRCRAFT SIMULATION. Analyses conducted revealed that only two basic parameters of flight need be controlled during PAR: heading and vertical speed. Thus, sophisticated flight equations need not be developed to implement heading and power controls for the demonstration. Simple transfer functions can be used. However, these transfer functions must be variable to reflect each of the three categories of aircraft identified earlier. Similar transfer functions can be used for wind and pilot factors.

PAR DISPLAY. For demonstration, a general-purpose CRT system can be used to simulate the display of the GCA CPN-4 system. The PAR display is actually two separate radar presentations displayed on one scope. The upper portion is the elevation display (EL) while the lower is the azimuth display (AZ). The scans are 7 degrees in elevation and 20 degrees in azimuth. The 'target' or 'pip' is about 0.50-inch high in the elevation display and 0.25-inch wide in the azimuth display. Based on information from the Navy GCA Syllabus, the simulated final approach leg should be displayed with a range of 50,000 feet or approximately 9.4 miles in length. The glide slope can be set at a 2.5-degree angle producing a glidepath intercept altitude of about 2180 feet. These values are considered optimal. The azimuth display has

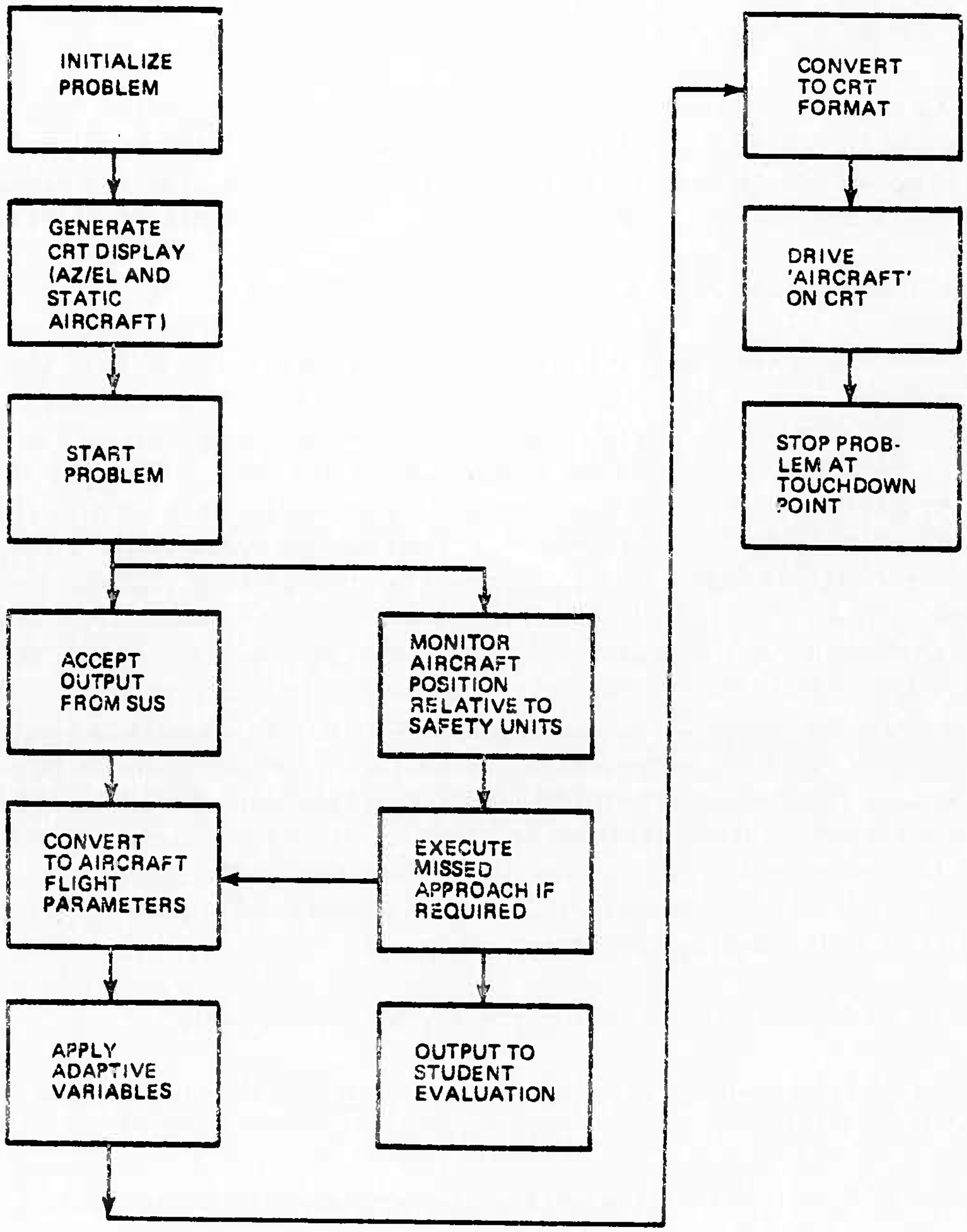


Figure 9. Training Control Block Diagram

the same range scale as the elevation display. Both displays should be logarithmic in range. Decision height will be 100 feet and the altitude over landing threshold 50 feet. These points occur at 0.66 and 0.33 mile, respectively, on the azimuth display. Safety limits for the elevation display should be included and can be based on flat terrain. An example of such a display is shown in figure 10.

For the demonstration, the 'target' should begin at a point 10 miles from touchdown and on runway heading. An altitude of 2180 feet can be maintained until intercept of the glidepath has occurred. Once intercept has occurred, 'target' movement will be a function of glidepath and course advisories.

TASK 14 – DESIGN SPEECH UNDERSTANDING SUBSYSTEM

The speech understanding subsystem (SUS) is the crux of the GCA controller training system. When operating properly, the SUS will accept one of approximately 80 advisories every 5 seconds or less, recognize it, and convert it to a functionally acceptable output (understanding). The SUS may be divided into sections which perform specific functions, as shown in figure 11. The system accepts advisories from the student, converts them to digital form, and extracts features which are used in recognition. These feature strings are sent to the central processor where they are compared to stored 'training' phrases. The resulting correlation of the spoken phrase to all stored phrases is formed into an array, ordered, and compared to a threshold. The value(s) above threshold is then sent through the recognition assistance routine where it is compared to the expected value(s) that is computed based on actual aircraft position and the advisory which should be issued. If there is a match, a final decision is made as to the phrase content, and the result is sent to the training control system and to the student speech evaluation subsystem. If the expected value(s) does not match the recognized phrase, one of three events may have occurred:

a. The phrase was spoken incorrectly by the student.

b. The phrase was spoken correctly, but the SUS failed to recognize it (rejection).

c. The SUS recognized the phrase incorrectly (false recognition).

The seriousness of these occurrences and their resolution must await detail design and development effort. For example, the resolution of the three events will be solved if high system accuracy is achieved.

Figure 11 illustrates the organization of the SUS. An overall block diagram of the GCA controller training system is shown in figure 12.

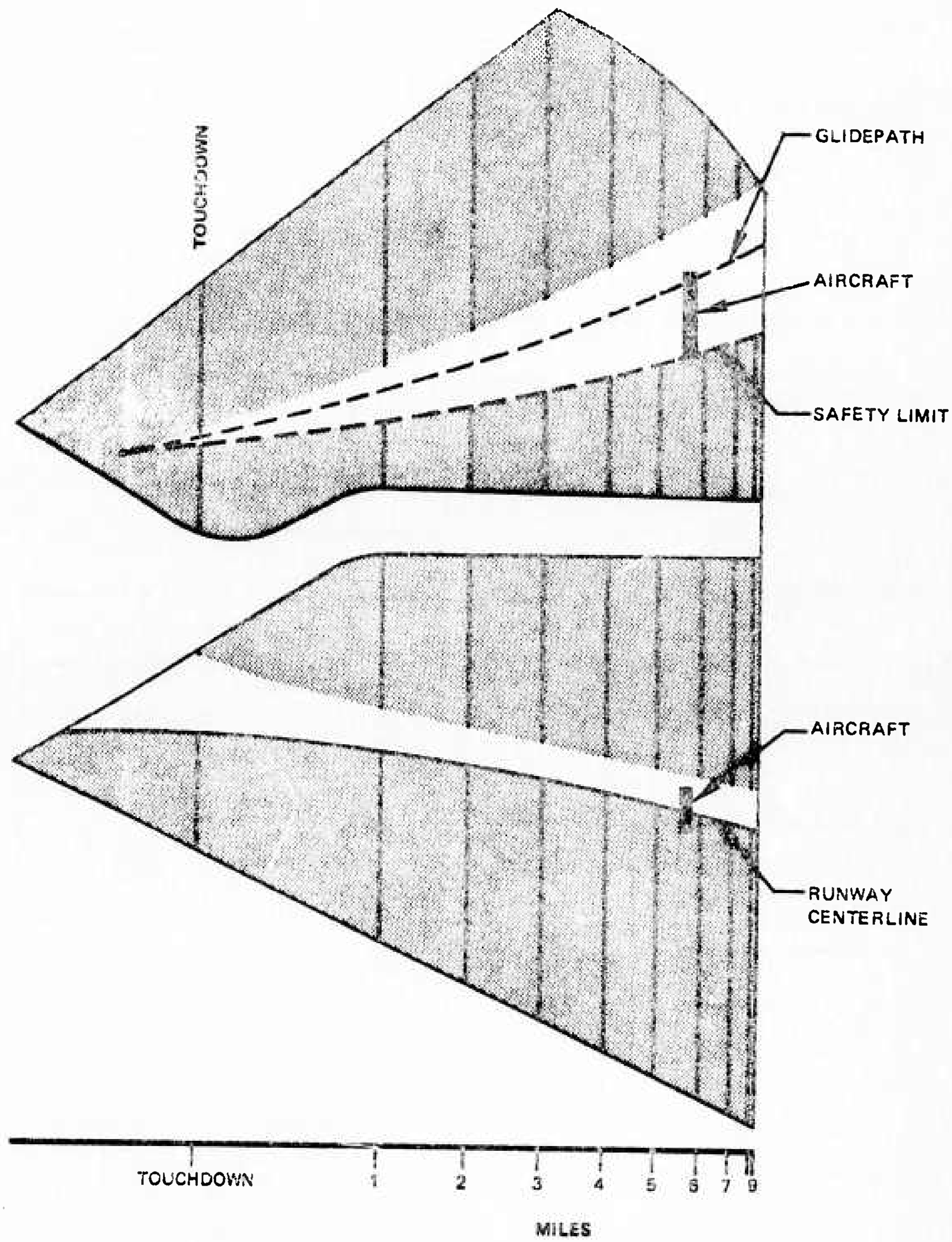


Figure 10. Typical PAR Display

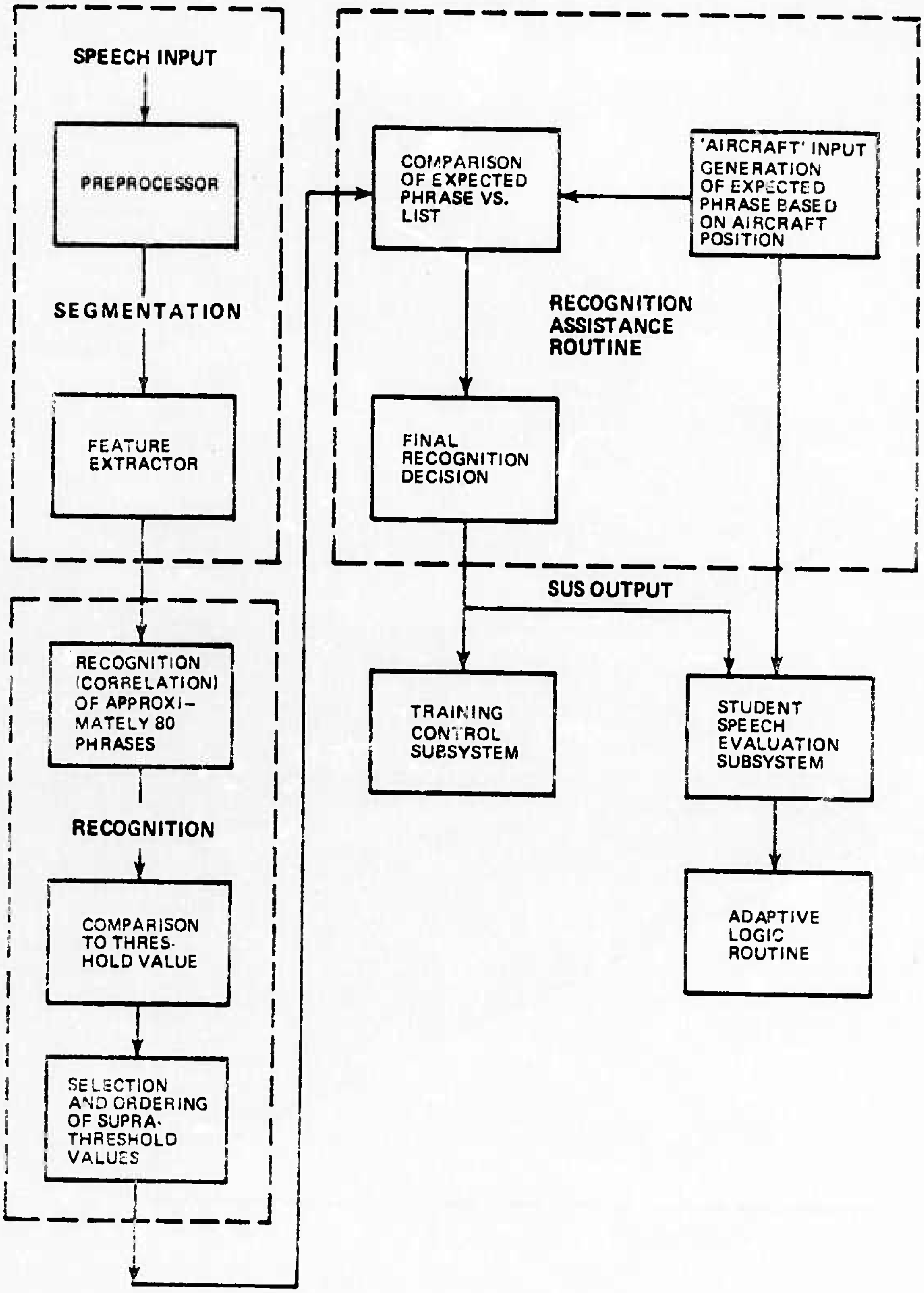


Figure 11. Speech Understanding Subsystem (SUS)

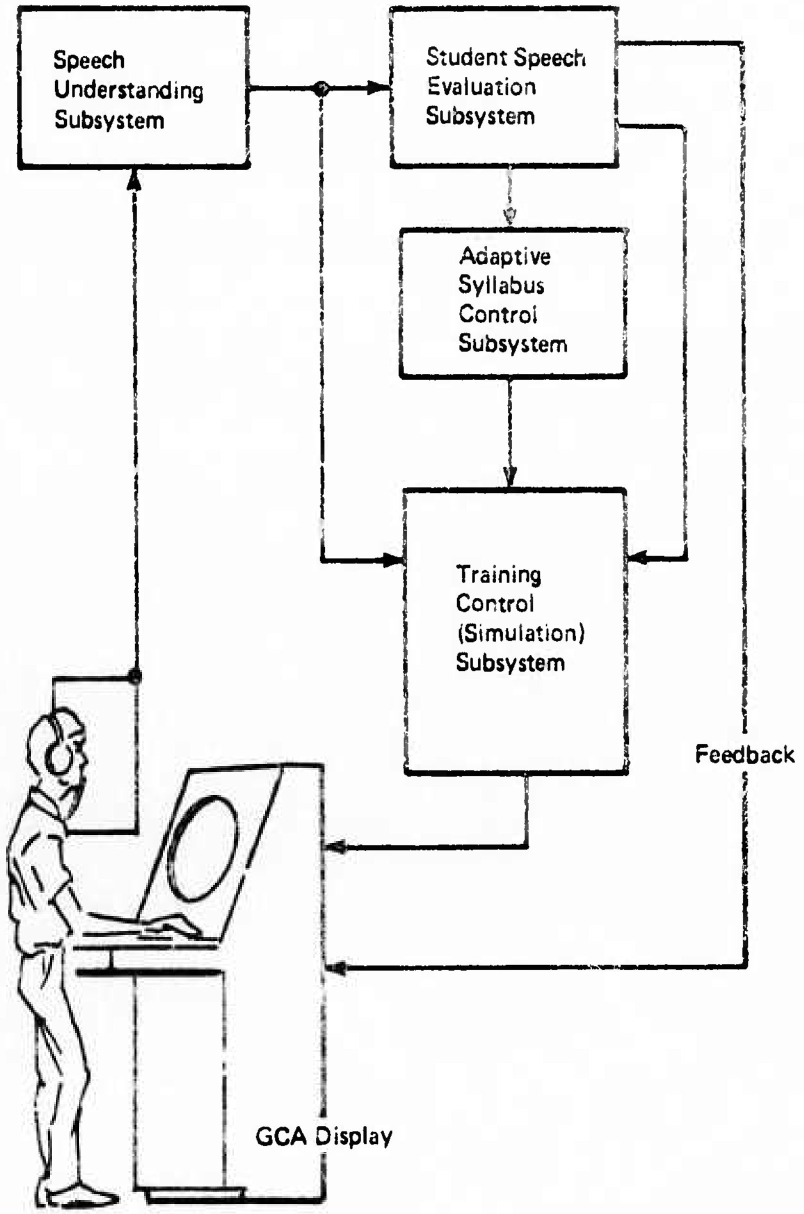


Figure 12. GCA Controller Training System Block Diagram

SECTION V

DISCUSSION

The review of the literature and the existing developments has revealed a rapidly expanding speech recognition and understanding technology. The effectiveness of this technology is magnified when the problem can be constrained by a small vocabulary, a rigid syntax, and a well-defined stimulus situation. The PAR phase of GCA operation is particularly amenable to automatic speech recognition in that controller phrases are short in length, finite in number, and regular in syntax. No minimal pairs were discovered, which significantly reduces the possibility of interword false recognition. Further developments should increase the observed recognition accuracy of 90 percent to 95 percent or greater.

There are still, however, several limitations on applications to training systems. For example, state-of-the-art speech understanding systems require the following:

a. Speech relatively free of 'uh's, er's, ah's'; elongated syllables such as well-ll-l-l-l; and unplanned pauses.

b. High similarity of speech between system training and system operation, particularly with respect to stress and intonation of the speaker.

c. Discrete pauses of at least 0.25 second between phrases.

d. System 'pretraining' for each new speaker.

e. Supplementary software routines to distinguish machine errors from student errors.

The first four limitations present little problem for experienced controllers since they have been trained to produce a consistent speech pattern with a definite rhythm. For a feasibility demonstration, however, untrained subjects may be used. This would require PAR vocabulary training to provide adequate knowledge of standard phrases. A prompting routine to facilitate the 'student task' during the demonstration might be considered.

The problem of detecting the type of error and the resolution can probably be solved during detailed design. The errors of consequence are 1) improper phraseology by the student, 2) failure to recognize a correct phrase, and 3) error in phrase recognition. The first error can result in two outcomes. Ideally, the system will reject the phrase as incorrect. However,

if the correlation with a stored phrase exceeds threshold, the system will falsely identify the incorrect phrase and output data. The probability of this occurrence must be minimized for effective training.

The second error (i. e., failure to recognize a correct phrase) is less serious although scoring will be affected. The outcome in terms of display change will be inconsequential and may well add realism to the simulation.

The third error, incorrectly identifying a correct phrase, can be serious. Fortunately, the GCA phraseology is highly redundant, and a great deal of information exists within the system to validate or verify the speech recognition output.

The redundancy within the phrases can reduce the speech to be identified within the advisory to one or two words. For example, the distinction between 'well above glidepath' and 'well below glidepath' reduces to a distinction between above and below. Moreover, machine errors, which stem from faulty threshold settings or matching errors, are expected to be qualitatively different from student errors resulting from confusion and faulty judgment. The true nature of these differences should be established via analyses conducted during the early phases of demonstration. Software routines can then be developed to resolve the problem based on other information available within the system.

SECTION VI

CONCLUSIONS AND RECOMMENDATIONS

The key issue in establishing the conceptual feasibility of an automated-adaptive training system for GCA controller training is whether or not automatic speech understanding technology has advanced sufficiently to meet the needs of such a system. In general, the study, after analyzing the requirements and technology and developing a feasible functional design, has reached an affirmative but qualified conclusion. The qualifications are not considered severe since the data developed indicate that system design solutions to the limitations can probably be achieved. The first qualification involves the speech understanding module itself; the second involves the development of supporting software routines to resolve ambiguity and error.

Existing and available speech understanding equipment is constrained to a limited number of 2-second samples of speech. The small number of samples problem can be solved by expanding the system computer memory. The patterns for the entire precision approach vocabulary, for example, can probably be handled by adding 8K of core to the processor. Expansion of the time sample to the required minimum 4 seconds (for PAR phraseology) was also investigated. A software change to achieve a 4-second sample has been successfully operated in the laboratory. Thus, it appears that the limitations (in terms of vocabulary and sample size) of existing equipment to meet the needs of a GCA controller training system can be technically solved.

The problem of ambiguities and errors in understanding or identifying GCA phrases poses the second qualification. The speech pattern criteria employed by all available speech recognition systems is not optimized for the GCA vocabulary. Filters and preprocessing units are not readily modified. However, additional software routines can be developed to compare or correlate on selected features to discriminate the GCA vocabulary as well as the wealth of redundant information available. Preliminary analyses and laboratory investigations have been successful in such modifications. However, detailed design and demonstration will be required before an evaluation to the GCA controller training system requirements can be made.

The system approach utilized in the study proved particularly effective in focusing attention on the major problems by isolating constraints and performance requirements early in the definition phase. These requirements or objectives provided the criteria for subsequent trade-off analyses of speech understanding equipment. Once the trade-off analyses had identified feasible equipment capability, conceptual design to the identified performance requirements was readily completed.

The design analyses isolated the development problems or risks. The problem areas which can only be solved by a prototype or breadboard design include:

a. Performance measurement.

b. Student speech variability.

c. Adaptive logic criteria.

d. Adaptive variable levels.

e. Student feedback and criteria.

f. Acoustic environment effects.

As can be seen, these risk or problem areas all revolve around the training system implementation and require resolution before an operational system can be developed and tested. Independent or isolated studies of the problems or factors would not solve the training system design problem since the problems are highly interactive. For example, ambient noise, student speech variability, performance measurement, and recognition thresholds will be highly interactive.

Therefore, although feasibility of an automated-adaptive GCA precision approach controller training system has been established, it is recommended that a laboratory version of the conceptual system be developed and implemented to explore and validate feasible solutions to the following problems or risk areas:

a. Recognition and discrimination of student's speech with the required accuracy for:

   1. Controlling simulation of aircraft and pilot during approach.

   2. Evaluation of student's control performance.

   3. Evaluation of student's knowledge of GCA procedures and vocabulary.

b. Processing of aircraft/pilot/controller data to:

   1. Validate and support automatic speech understanding.

   2. Generate syllabus difficulty factors.

   3. Establish the magnitude and effect of student speech variability, ambient noise, and related interactions.

c. Automated-adaptive training system integration; i. e., explore and solve system implementation problems before a field evaluation system is developed.

Thus, in summary, while conceptual feasibility is clearly indicated, development risks and missing design data require the development and exercising of a laboratory prototype system to explore the problems. The laboratory system should be limited to precision approach training (as the limiting condition), and simulation modules need only be general to explore the variables and factors outlined.

REFERENCES

Charles, J. P. and Johnson, R. M. Automated Training Evaluation (ATE). NAVTRADEVCEN-70-C-0132-1. January 1972. Naval Training Equipment Center. Orlando, Fla.

Charles, J. P., Johnson, R. M. and Swink, J. R. Automated Flight Training (AFT) GCI/CIC Air Attack. NAVTRAEQUIPCEN-72-C-0108-1. July 1973. Naval Training Equipment Center. Orlando, Fla.

BIBLIOGRAPHY

BOOKS:

Campanella, S. J. and Coulter, D. C. A Statistical Analog of Perception. Weiant Wathen-Dunn (Ed.). Models for the Perception of Speech and Visual Form. Cambridge, Mass.; MIT Press; 1967; 453-458.

Carterette, E. C. A Simple Linear Model for Vowel Perception. Weiant Wathen-Dunn (Ed.). Models for the Perception of Speech and Visual Form. Cambridge, Mass.; MIT Press; 1967; 418-427.

Clarke, F. R. A Detection Theory Approach to the Study of Speech Perception. Weiant Wathen-Dunn (Ed.). Models for the Perception of Speech and Visual Form. Cambridge, Mass; MIT Press; 1967; 458-461.

Clowes, M. B. Pictorial Relationships – A Syntactic Approach. B. Meltzer and D. Michie (Eds.). Machine Intelligence. Edinburgh, Scotland; Edinburgh University Press; 1969; 361-383.

Flanagan, J. L. Speech Analysis, Synthesis and Perception. Springer-Verlag. New York. Heidelberg. Berlin; 1972.

Hayes, D. G. (Ed.) Readings in Automatic Language Parsing. New York, N. Y.; American Elsevier; 1966.

Hemdal, J. F., and Hughes, G. W. A Feature-Based Computer Recognition Program for the Modeling of Vowel Perception. Weiant Wathen-Dunn (Ed.). Models for the Perception of Speech and Visual Form. Cambridge, Mass.; MIT Press; 1967; 440-453.

Hill, D. R. Automatic Speech Recognition – A Problem for Machine Intelligence. N. L. Collins and D. Michie. (Eds.). Machine Intelligence. Edinburgh, Scotland; Oliver and Boyd; 1967; 199-266.

Hill, D. R. and Wacker, E. B. ESOTerIC II – An Approach to Practical Voice Control. B. Meltzer and D. Michie (Eds.). Machine Intelligence. Edinburgh, Scotland; Edinburgh University Press; 1969; 463-493.

Lai, D. C. A Criterion for the Selection of Speech Features in Speech Recognition Based on Comparison of Experiements. Weiant Wathen-Dunn (Ed.). Models for the Perception of Speech and Visual Form. Cambridge, Mass.; MIT Press; 1967; 413-417.

Lindsay, R. K. Inferential Memory as the Basis of Machines which Understand Natural Language. E. A. Feigenbaum and J. Feldman (Eds.). Computers and Thought. New York, N. Y.; McGraw-Hill; 1963; 217-233.

Nelson, A. L., Herscher, M. B., Martin, T. B., Zadell, H. J. and Falter, J. W. Acoustic Recognition by Analog Feature – Abstraction Techniques. Weiant Wathen-Dunn (Ed.). Models for the Perception of Speech and Visual Form. Cambridge, Mass.; MIT Press; 1967; 428-440.

Ramo, S. Man-Machine Communications in the Coming Technological Society. R. S. Hollitch and B. Wittman (Eds.). Computer Applications. New York, N. Y.; The Macmillan Company; 1961; 45-54.

Stevens, K. N. and Halle, M. Remarks on Analysis-by-Synthesis and Distinctive Features. Weiant Wathen-Dunn (Ed.). Models for the Perception of Speech and Visual Form. Cambridge, Mass.; MIT Press; 1967; 88-102.

Weizenbaum, J. Contextual Understanding by Computers. P. A. Kolers and Murray Eden (Eds.). Recognizing Patterns. Cambridge, Mass.; MIT Press; 1968; 170-193.

PERIODICALS:

Alter, R. Utilization of Contextual Constraints in Automatic Speech Recognition, IEEE Trans. Audio Electroacoust., 1968, AU-16, 6-11.

Astrahan, M. M. Speech Analysis by Clustering or the Hypophoneme Method, IEEE Trans. Audio Electroacoust., 1970, AU-19, 113-114.

Atal, B. S. and Hanauer, S. L. Speech Analysis and Synthesis by Linear Prediction of the Speech Wave, J. Acoust. Soc. Am., 1971, 9, (3), 637-655.

Beckmann, P. Computerization of English, IEEE Spectrum, 1971, 8, 20-27.

Bell, C. G., Fujisaki, H., Heinz, J. M., Stevens, K. N. and House, A. S. Reduction of Speech Spectra by Analysis-by-Synthesis Techniques, J. Acoust. Soc. Am., 1961, 33, 1725-1736.

Beninghof, W. J., Jr. and Ross, M. J. Investigation of an Efficient Representation of Speech Spectra for Segmentation and Classification of Speech Sounds. IEEE Trans. Audio Electroacoust., 1970, AU-18, 33-42.

Bezdel, W. and Bridle, J. S. Speech Recognition Using Zero-Crossing Measurements and Sequence Information, Proc. Inst. Electrical Engineers, 1969, 116, 6-7-623.

Bobrow, D. G. and Klatt, D. H. A Limited Speech Recognition System, Fall Joint Computer Conference, 1968, 305-318.

Bogner, R. E. and Flanagan, J. L. Frequency Muitiplication of Speech Signals, IEEE Trans. Audio Electroacoust., 1969, AU-17, 202-208.

Broad, D. J. Basic Directions in Automatic Speech Recognition, Int. J. Man-Machine Studies, 1972, 4, 105-118.

Broad, D. J. and Fertig, R. H. Formant-Frequency Trajectories in Selected CVC – Syllable Nuclei, J. Acoust. Soc. Am., 1970, 47, (6), 1572-1582.

Campanella, S. J. and Phyfe, D. Application of a Model of the Analog Ear to Speech Signal Analysis, IEEE Trans. Audio Electroacoust., 1968, AU-16, 26-35.

Cannon, M. W., Jr. A Method of Analysis and Recognition for Voiced Vowels, IEEE Trans. Audio Electroacoust., p968, AU-16, 154-158.

Cardozo, B. L. and Ritsma, R. J. On the Perception of Imperfect Periodicity, IEEE Trans. Audio Electroacoust., 1968, AU-16, 159-164.

Carpenter, B. E. and Lavington, S. H. The Influence of Human Factors on the Performance of a Real-Time Speech Recognition System, J. Acoust. Soc. Am., 1973, 53, 42-45.

Chiba, S. Machines that "Hear" the Spoken Word, New Scientist, 1967, 37, 706-708.

Clapper, G. L. Machine Looks, Listens, Learns, Electronics, 1967, 40, 91-102.

Clapper, G. L. Automatic Word Recognition, IEEE Spectrum, 1971, 16, 57-69.

Clark, M. T. Word Recognition by Means of Orthogonal Functions, IEEE Trans. Audio Electroacoust., 1970, AU-18, 304-317.

Clarke, D. C. and Wall, R. E. An Economical Program for Limited Parsing of English, Fall Joint Computer Conference, 1965, 307-316.

Comer, D. J. The Use of Waveform Asymmetry to Identify Voiced Sounds, IEEE Trans. Audio Electroacoust., 1968, AU-16, 500-506.

Dammann, J. E. Application of Adaptive Threshold Elements to the Recognition of Acoustic-Phonetic States, J. Acoust. Soc. Am., 1965, 38, 213-223.

Das, S.K. Segmentation of Utterances of a Known Phrase Using Linear Threshold Techniques, IEEE Trans. Audio Electroacoust., 1972, AU-20, 142-154.

Das, S.K. and Mohn, W.S. A Scheme for Speech Processing in Automatic Speaker Verification, IEEE Trans. Audio Electroacoust., 1971, AU-19, 32-43.

David, E.E., Jr. Artificial Auditory Recognition in Telephony, IBM J. Res. Dev., 1958, 2, 294-309.

Davis, K.H., Biddulph, R. and Balashek, S. Automatic Recognition of Spoken Digits, J. Acoust. Soc. Am., 1952, 24, 269-279.

DeMori, R. Speech Analysis and Recognition by Computer Using Zero-Crossing Information, Acustica, 1971, 25, 269-279.

DeMori, R. A Descriptive Technique for Automatic Speech Recognition, IEEE Trans. Audio Electroacoust., 1973, AU-21, 89-100.

DeMori, R., Gilli, L. and Meo, A.R. A Flexible Real-Time Recognizer of Spoken Words for Man-Machine Communication, Int. J. Man-Machine Studies, 1970, 2, 317-326.

Denes, P.B. The Design and Operation of the Mechanical Speech Recognizer at University College, London, J. Brit. Inst. Radio Engr., 1959, 19, 219-234.

Denes, P.B. On-Line Computers for Speech Research, IEEE Trans. Audio Electroacoust., 1970, AU-18, 418-425.

Denes, P. and Mathews, M.V. Spoken Digit Recognition Using Time-Frequency Pattern Matching, J. Acoust. Soc. Am., 1960, 32, 1450-1455.

Denes, P.B. and Von Keller, T.G. Articulatory Segmentation for Automatic Recognition of Speech, Proc. 6th Int. Cong. Acoustics, 1968, B, 143-146.

Dixon, N.R. and Tappert, C.C. Toward Objective Phonetic Transcription – An On-Line Interactive Technique for Machine-Processed Speech Data, IEEE Trans. Man-Machine Systems, 1970, 11, (4), 202-210.

Dixon, N. R. and Tappert, C. C. Derivation of Phonetic Representation by Combining Steady-State and Transemic Classification in Automatic Recognition of Continuous Speech, Proc. Int. Conf. Speech Comm. Proc., 1972, 319-321.

Dolansky, L. Choice of Base Signals in Speech Signal Analysis, IRE Trans. Audio, 1960, AU-8, 221-229.

Dolansky, L. and Tjernlund, P. On Certain Irregularities of Voiced-Speech Waveforms, IEEE Trans. Audio Electroacoust., 1968, AU-16, 51-56.

Drucker, H. Speech Processing in a High Ambient Noise Environment, IEEE Trans. Audio Electroacoust., 1968, AU-16, 165-168.

Dudley, H. and Balashek, S. Automatic Recognition of Phonetic Patterns in Speech, J. Acoust. Soc. Am., 1958, 30, 721-732.

Elder, H. A. On the Feasibility of Voice Input to an On-Line Computer Processing System, Comm. ACM, 1970, 13, 339-346.

Ewing, G. D. and Taylor, J. F. Computer Recognition of Speech Using Zero-Crossing Information, IEEE Trans. Audio Electroacoust., 1969, AU-17, 37-40.

Falter, J. W. and Otten, K. W. Cybernetics and Speech Communications: A Survey of Russian Literature, IEEE Trans. Audio Electroacoust., 1967, AU-15, 27-36.

Fatehchand, R. Machine Recognition of Spoken Words, Advances in Computers, 1960, 1, 193-229.

Forgie, J. W. and Forgie, C. D. Results Obtained From a Vowel Recognition Computer Program, J. Acoust. Soc. Am., 1959, 31, 1480-1489.

Fourcin, A. J. Speech Source Inference, IEEE Trans. Audio Electroacoust., 1968, AU-16, 65-67.

Fry, D. B. Theoretical Aspects of Mechanical Speech Recognition, J. Brit. Inst. Radio Engr., 1959, 19, 211-218.

Fry, D. B. and Denes, P. The Solution of Some Fundamental Problems in Mechanical Speech Recognition, Language and Speech, 1958, 1, 35-38.

Fujimura, O. An Approximation to Voice Aperiodicity, IEEE Trans. Audio Electroacoust., 1968, AU-16, 68-72.

Fujisaki, H. and Kawashima, T. The Roles of Pitch and Higher Formants in the Perception of Vowels, IEEE Trans. Audio Electroacoust., 1968, AU-16, 73-77.

Gerstman, L. J. Classification of Self-Normalized Vowels, IEEE Trans. Audio Electroacoust., 1968, AU-16, 78-80.

Glenn, J. W. and Hitchcock, M. H. With a Speech Pattern Classifier, Computer Listens to its Master's Voice, Electronics, 1971, 84-89.

Goldberg, A. J. A Visual Feature Indicator for the Severely Hard of Hearing, IEEE Trans. Audio Electroacoust., 1972, AU-20.

Halle, M. and Stevens, K. Speech Recognition: A Model and a Program for Research, IRE Trans. Inf. Theory, 1962, 22, 155-159.

Heinz, J. M., Lindblom, B. E. F., and Lindquist, J. Ch. K-G. Patterns of Residual Masking for Sounds with Speech-Like Characteristics, IEEE Trans. Audio Electroacoust., 1968, AU-16, 107-111.

Hellwarth, G. A. and Jones, G. D. Automatic Conditioning of Speech Signals, IEEE Trans. Audio Electroacoust., 1968, AU-16, 169-179.

Hill, D. R. An ESOTerIC Approach to some Problems in Automatic Speech Recognition, Int. J. Man-Machine Studies, 1969, 1, 101-121.

Hill, D. R. Man-Machine Interaction Using Speech, Advances in Computers, 1971, 11, 165-230.

Hillix, W. A. Use of Two Non-Acoustic Measures in Computer Recognition of Spoken Digits, J. Acoust. Soc. Am., 1963, 35, 1978-1984.

Hillix, W. A., Fry, M. N. and Hershman, R. L. Computer Recognition of Spoken Digits Based on Six Non-Acoustic Measures, J. Acoust. Soc. Am., 1965, 38, 790-796.

Hoffman, H. S. Study of some Cues in the Perception of the Voiced Stop Consonants, J. Acoust. Soc. Am., 1958, 30, 1035-

House, A. S., Chairman, Session Q. Speech Analysis. Contributed Papers., J. Acoust. Soc. Am., 1960, 32, 1516-

Ichikawa, A., Nakano, Y. and Nakata, K. Evaluation of Various Parameter Sets in Spoken Digit Recognition, IEEE Trans. Audio Electroacoust., 1973, AU-21, 202-209.

Itahashi, S., Makino, S. and Kido, K. Discrete-Word Recognition Utilizing a Word Dictionary and Phonological Rules, IEEE Trans. Audio Electroacoust., 1973, AU-21, 239-248.

Ito, M. R. and Donaldson, R. W. Zero-Crossing Measurements for Analysis and Recognition of Speech Sounds, IEEE Trans. Audio Electroacoust., 1971, AU-19, 235-242.

Josselson, H. H. Automatic Translation of Languages since 1960: A Linguists' View, Advances in Computers, 1971, 11, 2-54.

Kaiser, J. F. The Digital Filter and Speech Communication, IEEE Trans. Audio Electroacoust., 1968, AU-16, 180-183.

Kalikow, D. N. and Swets, J. A. Experiments with Computer-Controlled Displays in Second-Language Learning, IEEE Trans. Audio Electroacoust., 1972, AU-20, 23-28.

King, J. H., Jr. and Tunis, C. J. Some Experiments in Spoken Word Recognition, IBM J. Res. Dev., 1966, 65-79.

Klatt, D. H. and Stevens, K. N. On the Automatic Recognition of Continuous Speech: Implications from a Spectrogram-Reading Experiment, IEEE Trans. Audio Electroacoust., 1973, AU-21, 210-217.

Klein, W., Plomp, R. and Pols, L. C. W. Vowel Spectra, Vowel Spaces, and Vowel Identification, J. Acoust. Soc. Am., 1970, 48, 999-1009.

Klein, S. and Simmons, R. F. A Computational Approach to Grammatical Coding of English Words, J. Assoc. Computing Machinery, 1965, 334-347.

Labanov, B. M. More About the Structure of the Speech Signal and the Principles of its Analysis, IEEE Trans. Audio Electroacoust., 1970, AU-18, 316-317.

Lea, W. A. Establishing the Value of Voice Communication with Computers, IEEE Trans. Audio Electroacoust., 1968, AU-16, 184-197.

Lea, W. A. An Approach to Syntactic Recognition Without Phonemics, IEEE Trans. Audio Electroacoust., 1973, AU-21, 249-258.

Lecours, M. and Sparkes, J. J. Adaptive Spectral Analysis for Speech-Sound Recognition, IEEE Trans. Audio Electroacoust., 1968, AU-16, 523-525.

Levitt, H. Acoustic Analysis of Deaf Speech Using Digital Processing Techniques, IEEE Trans. Audio Electroacoust., 1972, AU-20, 35-41.

Li, K-P, Hughes, G. W. and House, A. S. Correlation Characteristics and Demensionality of Speech Spectra, J. Acoust. Soc. Am., 1969, 46, 1019-1025.

Li, K-P., Hughes, G. W. and Snow, T. B. Segment Classification in Continuous Speech, IEEE Trans. Audio Electroacoust., 1973, AU-21, 50-57.

Liberman, A. M. The Grammars of Speech and Language, Cog. Psychol., 1970, 1, 301-303.

Lindgren, N. Machine Recognition of Human Language Part I: Automatic Speech Recognition, IEEE Spectrum, 1965, 2, (3), 114-136.

Lindgren, N. Machine Recognition of Human Language Part II: Automatic Speech Recognition, IEEE Spectrum, April 1965, 45-59.

Lindgren, N. Directions for Speech Research, IEEE Spectrum, 1968, 5, 83-

Maissis, A., Fievet, F. and Walrave, P. H. A New Analog-Digital Filtering Method for Real-Time Speech Recognition, IEEE Trans. Audio Electroacoust., 1970, AU-18, 385-388.

Makhoul, J. Speaker Adaptation in a Limited Speech Recognition System, IEEE Trans. Computers, 1971, 1057-1062.

Makhoul, J. Spectral Analysis of Speech by Linear Prediction, IEEE Trans. Audio Electroacoust., 1973, AU-21, 140-148.

Maksym, J. N. Real-Time Pitch Extraction by Adaptive Prediction of the Speech Waveform, IEEE Trans. Audio Electroacoust., 1973, AU-21, 149-154.

Manley, H. J. Analysis – Synthesis of Connected Speech in Terms of Orthogonalized Exponentially Damped Sinusoids, J. Acoust. Soc. Am., 1963, 35, 464-473.

Markel, J. D. FFT Pruning, IEEE Trans. Audio Electroacoust., 1971. AU-19, 305-311.

Markel, J. D. Digital Inverse Filtering – A New Tool for Formant Trajectory Estimation, IEEE Trans. Audio Electroacoust., 1972, AU-20, 129-137.

Markel, J. D. The SIFT Algorithm for Fundamental Frequency Estimation, IEEE Trans. Audio Electroacoust., 1972, AU-20, 367-377.

Markel, J. D. and Carey, B. Digital Voice Processing with a Wave Function Representation, Fall Joint Computer Conference, 1970, 387-397.

Markel, J. D. and Gray, A. H., Jr. On Autocorrelation Equations as Applied to Speech Analysis, IEEE Trans. Audio Electroacoust., 1973, AU-21, 69-79.

Marril, T. Automatic Recognition of Speech, IRE Trans. Human Factors Electronics, 1961, HFE-2, 35-38.

Mathews, M. V., Miller, J. E. and David, E. E. Jr. Pitch Synchronous Analysis of Voiced Sounds, J. Acoust. Soc. Am., 1961, 33, 179-186.

Mattingly, I. G. Speech Cues and Sign Stimuli, Am. Sci., 1972, 60, 327-337.

McCarthy, J., Earnest, L. D., Reddy, D. R. and Vicens, P. J. A Computer with Hands, Eyes and Ears, Fall Joint Computer Conference, 1968, 329-338.

Mermelstein, P. Computer-Generated Spectrogram Displays for On-Line Speech Research, IEEE Trans. Audio Electroacoust., 1971, AU-19, 44-47.

Miller, G. A. Decision Units in the Perception of Speech, IRE Trans. Inf. Theory, 1965, IT-8, 81-83.

Miller, J. C., Ross, P. W. and Wine, C. M. An Adaptive Speech Recognition System Operating in a Remote Time-Shared Computer Environment, IEEE Trans. Audio Electroacoust., 1970, AU-18, 26-31.

Nagy, G. Classification Algorithms in Pattern Recognition, IEEE Trans. Audio Electroacoust., 1968, AU-16 203-

Neely, R. B. and Reddy, R. D. Speech Recognition in the Presence of Noise, Proc. 7th Int. Cong. Acoust. (Budapest, Hungary), 1971, 3, 177-180.

Olson, H. F. Speech Processing Systems, IEEE Spectrum, February, 1964, 90-102.

Olson, H. F., Belar, H. and Rogers, E. S. Speech Processing Techniques and Applications, IEEE Trans. Audio Electroacoust., 1967, AU-15, 120-126.

O'Neill, L. A. The Representation of Continuous Speech with a Periodically Sampled Orthogonal Basis, IEEE Trans. Audio Electroacoust., 1969, AU-17, 14-21.

Oppenheim, A. V. Speech Spectrograms Using the Fast Fourier Transform, IEEE Spectrum, August, 1970, 57-62.

Oppenheim, A. V. and Schafner, R. W. Homomorphic Analysis of Speech, IEEE Trans. Audio Electroacoust., 1968, AU-16, 221-226.

Oppenheim, A. V. and Weinstein, C. J. Effects of Finite Register Length in Digital Filtering and Fast Fourier Transform, Proc. IEEE, August, 1972, 957-976.

Otten, K. W. Approaches to the Machine Recognition of Conversational Speech, Advances in Computers, 1971, 11, 127-163.

Peterson, E. Frequency Detection and Speech Formants, J. Acoust. Soc. of Am., 1951, 23, 668-674.

Petrick, S. R. and Griffiths, T. V. On the Relative Efficiencies of Context-Free Grammar Recognizers, Assoc. Computing Machinery, Communications, 1965, 8, 289-299.

Pickett, J. M. Some Applications of Speech Analysis to Communication Aids for the Deaf, IEEE Trans. Audio Electroacoust., 1969, AU-17, 283-289.

Pickett, J. M. Status of Speech-Analyzing Communication Aids for the Deaf, IEEE Trans. Audio Electroacoust., 1972, AU-20, 3-8.

Pierce, J. R. Men, Machines and Languages, IEEE Spectrum, 1968, 5, 44-49.

Pierce, J. R. Whither Speech Recognition, J. Acoust. Soc. Am., 1969, 46, 1049-1051.

Plomp, R., Pols, L. C. W. and Van de Geer, J. P. Dimensional Analysis of Vowel Spectra, J. Acoust. Soc. Am., 1967, 41, 707-712.

Pols, L. C. W. Real-Time Recognition of Spoken Words, IEEE Trans. Computers, 1971, C-20, 972-978.

Pols, L. C. W., Th. Van Der Kamp, L. J. and Plomp, R. Perceptual and Physical Space of Vowel Sounds, J. Acoust. Soc. Am., 1969, 46, 458-467.

Pols, L. C. W., Tromp, H. R. C. and Plomp, R. Frequency Analysis of Dutch Vowels from 50 Male Speakers, J. Acoust. Soc. Am., 1973, 53, 42-45.

Pruzansky and Mathews, M. Talker Recognition Procedure Based on Analysis of Variance. J. Acoust. Soc. Am., 1964, 36, 2041-

Purton, R. F. Speech Recognition Using Autocorrelation Analysis, IEEE Trans. Audio Electroacoust., 1968, AU-16, 235-239.

Reddy, D. R. Segmentation of Speech Sounds, J. Acoust. Soc. Am., 1966, 40, 307-312.

Reddy, D. R. Phoneme Grouping for Speech Recognition, J. Acoust. Soc. Am., 1966, 41, 1295-1300.

Reddy, D. R. Computer Recognition of Connected Speech, J. Acoust. Soc. Am., 1967, 42, 329-347.

Reddy, D. R. Segment-Synchronization Problems in Speech Recognition, J. Acoust. Soc. Am., 1969, 46, 69 (A).

Reddy, D. R. Speech Recognition: Prospects for the Seventies, Int. Fed. Inf. Processing. Proceedings, 1971, 71, I-5-

Reddy, D. R., Erman, L. D. and Neely, R. B. A Model and a System for Machine Recognition of Speech, IEEE Trans. Audio Electroacoust., 1973, AU-21, 229-238.

Reddy, D. R. and Robinson, A. E. Phoneme-to-Grapheme Translation of English, IEEE Trans. Audio Electroacoust., 1968, AU-16, 240-246.

Reddy, D. R. and Vicens, P. J. A Procedure for the Segmentation of Connected Speech, J. Audio Eng. Soc., 1968, 16, 404-411.

Rosenberg, A. E. A Computer-Controlled System for the Subjective Evaluation of Speech Samples, IEEE Trans. Audio Electroacoust., 1969, AU-17, 216-221.

Ross, P. W. A Limited-Vocabulary Adaptive Speech Recognition System, J. Audio Eng. Soc., 1967, 15, 415-418.

Sager, N. Syntactic Analysis of Natural Language, Advances in Computers, 1967, 8, 153-186.

Sakai, T. and Doshita, S. The Automatic Speech Recognition System for Conversational Sound, IEEE Trans. Elect. Computers, 1963, EC-12, 835-846.

Scarr, R. W. S. Zero-Crossings as a Means of Obtaining Spectral Information in Speech Analysis, IEEE Trans. Audio Electroacoust., 1968, AU-16, 247-255.

Scarr, R. W. S. Normalization and Adaptation of Speech Data for Automatic Speech Recognition, Int. J. Man-Machine Studies, 1970, 2, 41-59.

Schafer, R. W. A Survey of Digital Speech Processing Techniques, IEEE Trans. Audio Electroacoust., 1972, AU-20, 28-35.

Schafer, R. W. and Rabiner, L. R. System for Automatic Formant Analysis of Voiced Speech, J. Acoust. Soc. Am., 1970, 47, 634-648.

Schafer, R. W. and Rabiner, L. R. Design and Simulation of a Speech Analysis-Synthesis System Based on a Short-Time Fourier Analysis, IEEE Trans. Audio Electroacoust., 1973, AU-21, 165-174.

Schroeder, M. R. Parameter Estimation in Speech: A Lesson in Unorthodoxy, Proceedings of the IEEE, 1970, 58, 707-712.

Sebesteyen, G. Automatic Recognition of Spoken Numerals, J. Acoust. Soc. Am., 1960, 32, 1516-

Shearme, J. N. and Leach, P. F. Some Experiments with a Simple Word Recognition System. IEEE Trans. Audio Electroacoustics, 1967, AU-16, 256-261.

Shoenfelt, J. E., Tappert, C. C. and Goetze, A. J. Techniques for Efficient Encoding of Features in Pattern Recognition, IEEE Trans. Computers, 1971, 20, 1104-1106.

Simmons, R. F. Computational Linguistics. Answering English Questions by Computer: A Survey, Comm. ACM, 1965, 8, 53-70.

Simmons, R. F. Natural Language Question – Answering Systems: 1969, Comm. ACM, 1970, 13, 15-30.

Simmons, R. F., Burger, J. F. and Schwarcz, R. M. A Computational Model of Verbal Understanding, Fall Joint Computer Conference, 1968, 441-456.

Sitton, G. A. Acoustic Segmentation of Speech, Int. J. Man-Machine Studies, 1970, 2, 61-102.

Stevens, K. N. Toward a Model for Speech Recognition, J. Acoust. Soc. Am., 1960, 32, 47-55.

Tappert, C. C., Dixon, N. R. and Rabinowitz, A. S. Application of Sequential Decoding for Converting Phonetic to Graphic Representation in Automatic Recognition of Continuous Speech (ARCS), IEEE Trans. Audio Electroacoust., 1973, AU-21, 225-228.

Teacher, C. F., Kellet, H. G. and Focht, L. R. Experimental, Limited-Vocabulary Speech Recognizer, IEEE Trans. Audio Electroacoust., 1967, AU-15, 127-130.

Uhr, L. and Vossler, C. Recognition of Speech by a Computer Program That Was Written to Simulate a Model for Human Visual Pattern Perception, J. Acoust. Soc. Am., 1961, 33, 1426-

Vicens, P. J. Use of Syntax in the Analysis of Connected Speech, J. Acoust. Soc. Am., 1969, 46, 89(A).

Voiers, W. D. The Present State of Digital Vocoding Techniques: A Diagnostic Evaluation, IEEE Trans. Audio Electroacoust., 1968, AU-16. 275-279.

Von Keller, T. G. An On-Line Recognition System for Spoken Digits, J. Acoust. Soc. Am., 1971, 49, 1288-1295.

Warren, J. H. A Pattern Classification Technique for Speech Recognition, IEEE Trans. Audio Electroacoust., 1971, AU-19, 281-285.

Weitzman, C. Voice Recognition and Response Systems, IEEE Spectrum, December, 1969, 165-170.

Weizenbaum, J. ELIZA - a Computer Program for the Study of Natural Language Communications between Man and Machine, Comm. ACM, 1966, 9, 36-45.

Willemain, T. R. Tactile Pitch Displays for the Deaf, IEEE Trans. Audio Electroacoust., 1972, AU-20.

Wilson, Q. C. Optical Processing in Speech Recognizer Development, IEEE Trans. Audio Electroacoustics, 1968, AU-16, 280-287.

Wiren, J. and Stubbs, H. L. Electronic Binary Selection System for Phoneme Classification, J. Acoust. Soc. Am., 1956, 28, 1081-1091.

Woods, W. A. Transition Network Grammars for Natural Language Analysis, Comm. ACM, 1970, 13, 591-606.

REPORTS:

Babcock, M. L. Asynchronous Sampling of Speech with Some Vocoder Experimental Results. Final Report. Biological Computer Laboratory. University of Illinois. Urbana, Ill.

Bobrow, D. G., Hartley, A. K. and Klatt, D. H. A Limited Speech Recognition System II. Report No. 1819. 1969. Bolt, Beranek and Newman. Cambridge, Mass.

Broad, D. J. Formants in Automatic Speech Recognition. Interim Scientific Report. May, 1970. Speech Communications Research Laboratory. Santa Barbara, Calif.

Cooper, F. S. et al. Speech Research: A Report on the Status and Progress of Studies. SR-27. Oct., 1971. Haskins Laboratories, New Haven, Conn.

Cooper, F. S. et al. Speech Research: A Report on the Status and Progress of Studies. SR-29/30. July, 1972. Haskins Laboratories. New Haven, Conn.

Dersch, W. C. Decision Logic for Speech Recognition. Technical Report 16.01.106.018. 1961. IBM. San Jose, Calif.

Feuge, et al. A Review of Selected Articles in Automatic Speech Recognition. Report. Sept., 1973. Logicon, Inc. San Diego, Calif.

Gold, B. Word Recognition Computer Program. Technical Report 452. 1966. Lincoln Laboratories. MIT. Cambridge, Mass.

Heidorn, G. E. Natural Language Inputs to a Simulation Programming System – An Introduction. Technical Report NR-049-314. December, 1971. Navy Postgraduate School. Monterey, Calif.

Hughes, G. W. The Recognition of Speech by Machine. Technical Report 395. 1961. Laboratory of Electronics. MIT. Cambridge, Mass.

Hughes, G. W., House, A. S. and Kung-Pu, L. Research on Word Spotting. Final Technical Report. May, 1969. Air Force Cambridge Research Laboratories. Bedford, Mass.

Hunt, G. L. Automatic Recognition of Spoken English Words. Air Force Office of Scientific Research Report. July, 1972. Texas University. Austin, Texas.

Hyde, S. R. Automatic Speech Recognition: Literature Survey and Discussion. Report No. 45. 1969. Post Office Research Department. Dollis Hill, London, England.

Jakobson, R., Fant, C. G. M. and Halle, M. Preliminaries to Speech Analysis. Technical Report No. 13. 1952. Acoustical Laboratory, MIT. Cambridge, Mass.

Lea, W. A., Medress, M. F. and Skinner, T. E. Prosodic Aids to Speech Recognition. Report No. PX7940. October, 1972. Advanced Research Projects Agency. Arlington, Va.

Lehman, W. R., Stachowitz, R. A. et al. Feasibility Study on Fully Automatic High Quality Transmission. Report No. RADC-TR-71-295. Linguisitcs Research Center. University of Texas. Austin, Texas.

Lehmann, W. P. and Stachowitz, R. A. Normalization of Natural Language for Information Retrieval. Final Technical Report. Linguistics Research Center. University of Texas. Austin, Texas.

Makhoul, J. Speaker-Machine Interaction in Automatic Speech Recognition. Technical Report 480. Dec., 1970. MIT. Cambridge, Mass.

Makhoul, J. I. and Wolf, J. J. Linear Prediction and the Spectral Analysis of Speech. Report No. 2304. Bolt, Beranek and Newman. Cambridge, Mass.

Markel, J. D. Formant Trajectory Estimation from a Linear Least-Squares Inverse Filter Formulation. Oct., 1971. Speech Communications Research Laboratory. Santa Barbara, Calif.

Martin, T. B., Nelson, A. L. and Zadell, H. J. Speech Recognition by Feature Abstraction Techniques. USAF Technical Documentation Report No. AL-TDR-64-176. 1964.

Martin, T. B. and Talvage, J. J. Speech Pattern Recognition by Simulated Neural Networks. ASD-TDR No. 62-511. 1962. USAF Aerospace Systems Division. W-P AFB, Ohio.

Martin, T. B., Zadell, H. J., Nelson, A. L. and Cox, R. B. Recognition of Continuous Speech by Feature Abstraction. TR-66-189. 1966. RCA Advanced Technology Laboratory. Camden, N. J.

Newell, A. et al. Speech Understanding Systems: Final Report of a Study Group. May, 1971. Carnegie-Mellon University. Pittsburgh, Pa.

Paulus, E. Bibliography on Automatic Speech Recognition. TR 25.064. 1966. IBM Laboratory. Vienna, Austria.

Pezoldt, V. J. and Persensky, J. J. Human Factors Evaluation of a Speech Understanding System. Final Report. March, 1973. National Bureau of Standards. Washington, D. C.

Preusse, J. W. Word Recognition and Speaker Authentication Using Amplitude Independent and Time Independent Word Features. R&DTR ECOM-3439. June, 1971. U.S. Army Electronics Command. Fort Monmouth, N. J.

Reddy, D. R. An Approach to Computer Speech Recognition by Direct Analysis of the Speech Wave. TR CS49. 1966. Stanford University. Stanford, Calif.

Richter, W. J. et al. Investigation of Speech Pattern-Matching. Final Technical Report. Sept., 1971. MELPAR. Falls Church, Va.

Stevens, K. N. Study of Acoustic Properties of Speech Sounds II, and Some Remarks on the Use of Acoustic Data in Schemes for Machine Recognition of Speech. August, 1969. Bolt, Beranek and Newman, Inc. Cambridge, Mass.

Stevens, K. N. and Klatt, M. M. Study of Acoustic Properties of Speech Sounds. S. R. No. 8. August, 1968. Bolt, Beranek and Newman, Inc. Cambridge, Mass.

Talbert, L. R. et al. A Real-Time Adaptive Speech Recognition System. TR No. 6760. May, 1963. Stanford Research Electronics Laboratory. Stanford, Calif.

Teacher, C. F. and Piotrowski, C. Voiced Sound Recognition. RADC-TR-65-184. July, 1965. Rome Air Development Center. Rome, N. Y.

Weiss, M. R., Parsons, T. W. and Aschkenasy, E. Implementation of an On-Line System for Automatic Recognition of Spoken Words. Final Technical Report. December, 1971. Federal Scientific Corporation. New York, N. Y.

Yilmaz, H. et al. A Real-Time, Small Vocabulary, Connected-Word Speech Recognition System. Final Technical Report. November, 1972. Perception Technology Corporation. Winchester, Maine.

______. Speech Understanding Systems. Semiannual Technical Summary Report. December, 1972 – May, 1973. Advanced Research Projects Agency. Lincoln Laboratories. Cambridge, Mass.

OTHER:

Bryden, B. Speech Recognition by Machine. M. Sc. Thesis. 1970. Department of Electrical Engineering. University of Calgary. Alberta, Canada.

Evans, T. G. A Grammar-Based Pattern Analyzing Procedure. 1968. Proceedings of the IFIP Conference. Edinburgh, Scotland.

Frick, F. C. Research on Speech Recognition at Lincoln Laboratory. 1959. Proceedings of the Seminar on Speech Compression and Coding. AFCRL PB 146624. L. G. Hanscom Field, Bedford, Mass.

Hill, D. R. An Abstracted Bibliography of Some Papers in Automatic Speech Recognition. 1967. Technical Memo 522. Standard Telecommunications Laboratories. Harlow, England.

Itakura, F. and Saito, S. Speech Analysis-Synthesis System Based on Autocorrelation Coefficients. 1969. Meeting of the Acoustical Society of Japan.

Klatt, D. H. and Stevens, K. N. Sentence Recognition from Visual Examination of Spectrograms and Machine-Aided Lexical Searching. 1972. Proceedings of the International Conference on Speech Communications Processing. 315-318.

Lea, W. A. Ultimate Goals for Effective Man-Computer Speech Communication. August 1968. U. S. G. Memorandum.

Lindblom, B. and Svensson, S-G. Interaction Between Segmental and Non-Segmental Factors in Speech Recognition. 1972. Proceedings of the International Conference on Speech Communications Processing. 331-333.

Martin, T. B. Acoustic Recognition of a Limited Vocabulary in Continuous Speech. 1970. Doctoral Dissertation. University of Pennsylvania. Philadelphia, Pa.

Medress, M. A. A Procedure for the Machine Recognition of Speech. 1972. Proceedings of the International Conference on Speech Communications Processing. 113-116.

Meeker, W. F. Parametric Recognition of Speech Sounds. 1959. Proceedings of the Seminar on Speech Compression Coding. PB 146624. AFCRL. LG Hanscom Field. Bedford, Mass.

Merserean, R. M. The Digital Reconstruction of Multidimensional Signals from Their Projections. October, 1972. Paper Presented at the Tenth Alberton Conference on Circuit and System Theory. Monticello, Ill.

Otten, K. W. The Formant Vocoder and its Use for Automatic Speech Recognition. 1961. Proceedings of the 3rd International Conference on Acoustics. Elsevier, Amsterdam.

Purton, R. F. An Automatic Word Recognizer Based on Autocorrelation Analysis. 1968. Reprint from Institute of Electrical Engineering. London, England.

Reddy, D. R., Erman, L. D. and Neely, R. B. The C-MU Speech Recognition Project. 1970. Proceedings of the IEEE Systems Science Cybernetics Conference.

Sergeant, R. L. and Murry, T. Processing Helium Speech. May, 1972. Naval Submarine Medical Research Laboratory.

Shoup, J. E. Research on Communications and Automatic Speech Recognition. 1970. Speech Communications Research Laboratory. Santa Barbara, Calif.

Vicens, P. Aspects of Speech Recognition by Computer. 1969. Doctoral Dissertation. Stanford University. Stanford, Calif.

Von Keller, T. G. Automatic Recognition of Spoken Words. 1968. 75th Meeting of the Acoustical Society of America. Ottawa, Canada.

Wakita, H. Estimation of the Vocal Tract Shape by Optimal Inverse Filtering and Acoustic/Articulatory Conversion Methods. 1972. Speech Research Communications Laboratory. Santa Barbara, Calif.

Winograd, T. Procedures as a Representation of Knowledge in a Computer Program for Understanding Natural Language. 1971. Artificial Intelligence Memo. MIT. Cambridge, Mass.

______. Conference Reprints: Conference on Speech Communication and Processing. November, 1967 AFCRL. Hanscom Field. Bedford, Mass.

APPENDIX A

PAR PHRASE LIST WITH CRITERIA

TABLE 4. PAR PHRASE LIST WITH CRITERIA

Regular PAR Approach	Criteria for Issuance
Approaching glidepath	top of target touches glidepath cursor (10–30 sec prior to intercept)
Begin descent	target intercepts glidepath cursor
On glidepath	target bisected by glidepath cursor
Going above glidepath	center of target leaves the glidepath cursor (upward)
Going rapidly above glidepath	center of target leaves the glidepath cursor (upward), rapidly
Going slowly above glidepath	center of target leaves the glidepath cursor (upward), slowly
Slightly above glidepath	2/3's of target above glidepath cursor
Slightly above glidepath and holding	2/3's of target above glidepath cursor, unchanged
Slightly above glidepath and coming down	2/3's of target above glidepath cursor, with movement toward cursor
Slightly above glidepath and coming rapidly down	2/3's of target above glidepath cursor, with rapid movement toward cursor
Slightly above glidepath and coming slowly down	2/3's of target above glidepath cursor, with slow movement toward cursor
Slightly above glidepath and going further above	2/3's of target above glidepath cursor, movement from slightly above to above
Slightly above glidepath and going rapidly further above	2/3's of target above glidepath cursor, rapid movement from slightly above to above
Slightly above glidepath and going slowly further above	2/3's of target above glidepath cursor, slow movement from slightly above to above
Well above glidepath	bottom of target breaks contact with glidepath cursor
Well above glidepath and holding	bottom of target breaks contact with glidepath cursor, unchanged
Well above glidepath and coming down	bottom of target breaks contact with glidepath cursor, with movement toward cursor
Well above glidepath and coming rapidly down	bottom of target breaks contact with glidepath cursor, with rapid movement toward cursor
Well above glidepath and coming slowly down	bottom of target breaks contact with glidepath cursor, with slow movement toward cursor
Well above glidepath and going further above	bottom of target breaks contact with glidepath cursor, movement from well above to beyond
Well above glidepath and going rapidly further above	bottom of target breaks contact with glidepath cursor, rapid movement from well above to beyond
Well above glidepath and going slowly further above	bottom of target breaks contact with glidepath cursor, slow movement from well above to beyond
Going below glidepath	center of target leaves glidepath cursor (downward)
Going rapidly below glidepath	center of target leaves glidepath cursor (downward), rapidly
Going slowly below glidepath	center of target leaves glidepath cursor (downward), slowly
Slightly below glidepath	2/3's of target is below the glidepath cursor
Slightly below glidepath and holding	2/3's of target is below the glidepath cursor, unchanged
Slightly below glidepath and coming up	2/3's of target is below the glidepath cursor, movement toward cursor
Above glidepath	bottom of target touches glidepath
Below glidepath	top of target touches glidepath

*Target refers to radar return of controlled aircraft as shown on CRT

TABLE 4. PAR PHRASE LIST WITH CRITERIA (Cont)

Regular PAR Approach	Criteria for Issuance
Slightly below glidepath and coming rapidly up	2/3's of target below glidepath cursor, rapid movement toward the cursor
Slightly below glidepath and coming slowly up	2/3's of target below glidepath cursor, slow movement toward the cursor
Slightly below glidepath and going further below	2/3's of target below glidepath cursor, movement from slightly below to below
Slightly below glidepath and going rapidly further below	2/3's of target below glidepath cursor, rapid movement from slightly below to below
Slightly below glidepath and going slowly further below	2/3's of target below glidepath cursor, slow movement from slightly below to below
Well below glidepath	top of target breaks contact with glidepath cursor
Well below glidepath and holding	top of target breaks contact with glidepath cursor, unchanged
Well below glidepath and coming up	top of target breaks contact with glidepath cursor, movement toward cursor
Well below glidepath and coming rapidly up	top of target breaks contact with glidepath cursor, rapid movement toward cursor
Well below glidepath and coming slowly up	top of target breaks contact with glidepath cursor, slow movement toward cursor
Well below glidepath and going further below	top of target breaks contact with glidepath cursor, movement from well below to beyond
Well below glidepath and going rapidly further below	top of target breaks contact with glidepath cursor, rapid movement from well below to beyond
Well below glidepath and going slowly further below	top of target breaks contact with glidepath cursor, slow movement from well below to beyond
On course	target bisected by azimuth cursor
Turn right heading ______	target left of desired course
Turn left heading ______	target right of desired course
Heading is ______	establish heading
Going right of course	target center leaves the azimuth cursor, right
Going left of course	target center leaves the azimuth cursor, left
Going rapidly right of course	target center rapidly leaves the azimuth cursor, right
Going slowly right of course	target center slowly leaves the azimuth cursor, right
Going rapidly left of course	target center rapidly leaves the azimuth cursor, left
Going slowly left of course	target center slowly leaves the azimuth cursor, left
Right of course	left edge of cursor touches azimuth cursor
Left of course	right edge of cursor touches azimuth cursor
Slightly right of course	2/3's of target is right of azimuth cursor
Slightly right of course and holding	2/3's of target is right of azimuth cursor, unchanging
Slightly right of course and correcting	2/3's of target is right of azimuth cursor, movement toward cursor
Slightly left of course	2/3's of target is left of azimuth cursor
Slightly left of course and holding	2/3's of target is left of azimuth cursor, unchanging
Slightly left of course and correcting	2/3's of target is left of azimuth cursor, movement toward cursor

TABLE 4. PAR PHRASE LIST WITH CRITERIA (Cont)

Regular PAR Approach	Criteria for Issuance
Right of course and holding	left edge of target touches azimuth cursor, unchanged
Right of course and correcting	left edge of target touches azimuth cursor, movement toward cursor
Left of course and holding	right edge of target touches azimuth cursor, unchanged
Left of course and correcting	right edge of target touches azimuth cursor, movement toward cursor
Well right of course	target breaks contact with azimuth cursor (right)
Well right of course and holding	target breaks contact with azimuth cursor (right), unchanged
Well right of course and correcting	target breaks contact with azimuth cursor (right, movement toward cursor
Well left of course	target breaks contact with azimuth cursor (left)
Well left of course and holding	target breaks contact with azimuth cursor (left), unchanged
Well left of course and correcting	target breaks contact with azimuth cursor (left), movement toward cursor
Wind is ___ ___ ___ at ___ ___.	surface wind
Cleared to land runway ___ ___ right.	tower clearance given, 3–4 miles from touchdown
Cleared to land runway ___ ___ left.	tower clearance given, 3–4 miles from touchdown
At decision height	2/3 miles from touchdown (published decision height)
No glidepath information available	loss of altit[illegible] information (glidepath)
___ miles from touchdown	range at point where leading edge to target touches mile marker
Over landing threshold, centerline is right/left	approximately 1/3 mile from touchdown
Over approach lights	approximately 1/2 mile from touchdown
Execute missed approach	a. safety limits are exceeded b. radical aircraft maneuver is observed c. position of aircraft is in doubt d. identification of aircraft is in doubt e. radar contact is lost f. bad airport conditions g. bad airport traffic

APPENDIX B

TIME-LINE ANALYSIS OF PAR CONTROLLER TASKS

TABLE 5. TYPICAL TIME SEQUENCE OF PAR CONTROLLER OUTPUT

Cumulated PAR Time (in seconds)	Controller Output
00	Approaching glidepath; begin descent flight
05	Below glidepath, coming up
07	Coming up and on glidepath
08	Six miles from touchdown
12	On glidepath, heading 233
14	On glidepath
21	On glidepath, heading 233; the centerline is left; correcting
25	On glidepath, 5 miles from touchdown
29	Going above, slightly above glidepath
32	Slightly above glidepath
37	Slightly above glidepath
39	Turn right, heading 235
42	Slightly above glidepath
50	Turn right, heading 238; 4 miles from touchdown
53	Going above glidepath
56	Turn right, heading 242
60	Above glidepath
63	Above glidepath; heading 242

TABLE 5. TYPICAL TIME SEQUENCE OF PAR CONTROLLER OUTPUT (Cont)

Cumulated PAR Time (in seconds)	Controller Output
68	Slightly above glidepath, coming down
72	Slightly above glidepath
75	Three miles from touchdown; transmission break
80	Coming down and on glidepath; heading 242, on course
87	On glidepath; 2-1/2 miles from touchdown; the wind is 350° at 14
90	Cleared to land, runway 24, right.
92	Now on glidepath
95	Turn left heading 240
97	Now on glidepath
102	Going slightly below glidepath, slightly below
105	Below glidepath
107	Going further below glidepath
109	Below glidepath and coming up
112	One one-half miles from touchdown; heading 245
114	Slightly below glidepath
120	Slightly below glidepath; turn right heading 248; the centerline is right
123	Coming up and on glidepath
126	One mile from touchdown; the centerline is right

TABLE 5. TYPICAL TIME SEQUENCE OF PAR CONTROLLER OUTPUT (Cont)

Cumulated PAR Time (in seconds)	Controller Output
132	Turn right heading 250; on glidepath (contiguous with above phrase)
138	On glidepath; 1/2 mile; centerline is right, correcting
141	On glidepath
144	On glidepath; heading 248; on glidepath
149	Over landing threshold
150	Over touchdown
155	Squawk, standby, contact tower at 315.6
Total time = 2 minutes, 35 seconds.	

APPENDIX C

REVIEW, SELECTION, AND TEST OF SPEECH RECOGNITION DEVICES

REVIEW OF SPEECH RECOGNITION DEVICES

A review of the literature in automatic speech recognition (ASR) between 1950 and the present revealed at least 20 attempts to build devices that recognize and/or understand human speech. Although the majority of these devices have been built in the United States, others have been attempted in the U. S. S. R., Great Britain, and the Scandanavian countries. Of those in the United States, the majority have been built by large corporations and smaller research firms attached to universities. Supporting research funding has been supplied by the Advanced Research Projects Agency (ARPA) of the U. S. Government. The funding for devices has come from a variety of sources, including corporate development, government (Air Force, Post Office), and university funds.

The success of the approaches varies with the scope and approach taken by the various developers. The scope of the development determined whether the intended application required 'isolated word recognition' as opposed to 'continuous speech' or 'connected word' speech and determined whether the application required understanding in addition to recognition. Finally, the approach determined whether digital or analog techniques were used, whether pattern-matching or segmentation was employed, and whether the recognition element was phonemic or feature oriented. Of these approaches, most were directed to the isolated or connected word problem with small vocabularies, and used the analog segmentation approach. Only four of the devices are known to be commercially available today. These are made by:

a. Scope Electronics.

b. Threshold Technology, Inc.

c. Perception Technology, Inc.

d. Culler-Harrison, Inc.

The first three companies were visited and their system discussed to determine its capability. The fourth company presently produces only a sound analysis device, but it can be used for speech with some modification. Telephone discussions with their personnel indicated that work is in progress on a speech recognition device, but it has not been completed or tested.

The following is a description of the three devices, based on information from the visits, brochures, and reports.

a. Threshold Technology, Inc. – Threshold Technology, Inc. (TTI) builds a basic system named the VIP-100, which recognizes connected word phrases or isolated words from a limited

vocabulary (32 words or phrases) at 95-percent accuracy. The length of phrases is limited to 2.0 seconds and the vocabulary limited to 32 such phrases. These aspects of the VIP-100 are determined by software and core memory available and are, therefore, expandable.

In terms of configuration, the VIP-100 is composed of a special preprocessor (including an analog-to-digital converter) and feature extractor. The equipment includes a minicomputer (Nova 1200), an ASR 33 Teletypewriter, and a small display/control device for system operation. The interface between the Nova 1200 and the preprocessor/feature extractor unit is unique but employs standard input/output card connectors for the Nova 1200.

During operation, the VIP-100 operates on speech by segmenting each 2.0-second utterance into a matrix containing 512 bits of information (32 features mapped onto 16 time segments) when the end of the utterance is detected. The feature set contains five broad class features (vowel/vowel-like, long pause, short pause, unvoiced noise-like consonants, and bursts) and 27 phoneme-like features. Of these 27 phoneme-like features, 15 are vowel indicants, thus stressing the importance of vowel detection. These features are derived by forcing each utterance through a set of 19 bandpass filters, ranging in center frequency from 260 Hz to 7626 Hz. The output of these filters is full-wave rectified and logarithmically compressed to provide a 50 db dynamic range from which ratio measurement is possible. From various combinations and sequences of these outputs, a significant set of features is derived which is the spectral derivative (de/df) indicative of the overall spectrum shape. Measurement of the slopes of spectrum energy changes ensures detection of peaks which, in turn, serve to identify formants. Formants (energy concentrations at particular frequencies) are known to be crucial to vowel detection and discrimination.

Recognition is performed in the Nova 1200 by comparison of real-time generated matrices to stored matrices from a training session. To recognize phrases or words, the VIP-100 requires that each utterance be spoken 10 consecutive times. This allows the VIP-100 to develop a consistent matrix against which utterances may be digitally compared for recognition. Similarities and dissimilarities in each comparison are appropriately weighted, and the net result provides a weighted correlation product. Correlation products are also calculated after the input matrix has been shifted ± 1 time segment. The stored referent producing the

highest correlation when all stored matrices are compared to the input is selected as the 'recognition' match. To prevent the VIP-100 from being false-alarm prone, a threshold has been programmed which, in effect, causes the device to reject the utterance if some threshold correlation value is not met or exceeded.

Developments under way include improvement of the feature set, increase in recognition resolution by combining the number of samples taken, modifying the recognition technique to decrease dependence on speaker stress and intonation, and improving 'training' techniques.

b. Scope, Inc. – Scope, Inc. builds a basic system, named the VCS, which also recognizes connected word or isolated word speech from a limited vocabulary (24 phrases, each 1.0 second in length) with 90-percent accuracy. As with the VIP-100 device, phrase length and vocabulary are a function of software and memory, respectively.

In terms of configuration, the VCS consists of a spectrum analyzer, an analog-to-digital converter, a hard-wired processor (digital), memory, and an output register device. A standard minicomputer could be substituted for the digital processor and memory.

During operation, the spectrum analyzer divides each utterance into 16 frequency bands between 200 and 5000 Hz. The bands compose a power spectrum which is a frequency x amplitude x time representation of the speech signal. Samples from these bands are taken every 1/60 second and multiplexed onto a single channel where they are converted to digital form. Thus, the original utterance arrives at the processor as a string of four-bit binary numbers, each representing the amplitude of one of the 16 frequency bands at some instant in time. In the processor, these strings are compressed (normalized) into a fixed length (120 bit) code (pattern) which represents the salient features of the uttered spectrum. Thus, there is no segmentation in the VCS. Formants are detected by peaks in the spectral pattern.

During recognition, patterns generated by the real-time normalization routine are compared to reference patterns generated during training. During 'training,' five voicings of each word or phrase of the vocabulary are compressed into 120-bit patterns and stored in core memory. These 120 bits contain both salient and not-so-salient variations of the five utterances. These stored patterns are compared to the input utterance where they are matched bit by bit, summed, and compared to a threshold value.

The highest output of summed matched bits is accepted as the recognition if it is above threshold; otherwise, the VCS rejects the utterance as unfamiliar.

c. Perception Technology, Inc. – Perception Technology, Inc. (PTI) builds a basic adaptive speech recognition system which presently recognizes isolated, single-syllable digits or connected strings of digits up to 2.4 seconds in length with accuracies between 90 and 98 percent. A positive feature of the PTI device is its minimal speaker dependence which nearly eliminates the need for prior 'training.' If needed, training consists of voicing six key words three times each. This allows the adaption routine to make transformations which shift the speaker's voice characteristics toward the pre-established norm. This procedure accounts for accent, sex, and intonation differences observed between speakers.

In theory, the PTI concept differs markedly from other manufacturers, who tend to emphasize 'formant theory.' The PTI approach places speech theory in a class with relativity and color perception, stressing its polar and relativity features. The theory is based on the assumption that speech, like color, may be placed on a 'wheel' in which two dimensions (frequency and amplitude) are represented. Of specific interest in the theory is pitch (frequency) and loudness (amplitude). The two-dimensional speech wheel is divided into patches. These patches represent frequency combinations of utterances in a closed space. Once the space is divided for a certain vocabulary, the locus (trajectory) of points corresponding to words is analyzed for patch 'hits.' The identity of the patches entered by the trajectory of a word are processed by the decision algorithm in conjunction with 24 other features such as initial or final 'S,' $t_{5,9}$, etc. Word boundaries are determined by pauses in a 400-700 Hz bandpass filter or suitable dips in the energy spectrum. These 24 features are derived from the output of six bandpass filters between 250 and 5300 Hz.

The system is composed of a PDP-8E minicomputer with 8K of memory, an ASR-33 Teletypewriter, and a display unit.

While the PTI device is not sufficiently developed for the present application, planned development will bring it closer. Plans have been made by PTI to investigate the application of this theory to larger vocabularies.

SELECTION OF A SPEECH RECOGNITION DEVICE FOR THE SUS

Based primarily upon considerations of the capability of the device to perform the functions described in this report, the VIP-100 was selected as the tentative front-end device for the SUS. The selection of the VIP-100 was bolstered by new developments and improvements being undertaken by Threshold Technology, Inc. The VCS (Scope, Inc.) would require too extensive a modification to perform the subject tasks. The PTI device is not yet available.

VIP-100 TESTS

A test using selected PAR advisories was devised to check the recognition accuracy of the VIP-100 system. The test consisted of nine lists of 10 items each since the VIP-100 system used had storage for only a vocabulary of 10 items of 2 seconds or less in length. Five of the nine lists consisted of phrases, while the remaining four lists were single words. (Refer to table 6.) The first list included glidepath advisories; the second list, trend advisories; the third list, course corrections; the fourth list, headings (numerics); and the fifth list had information advisories.

The VIP-100 was trained on each list by repeating each item 10 times in succession (100 total repetitions). Once the device was trained on a given list, the items in that list were then randomly selected and spoken until all items had been repeated four times. This procedure tested for intralist false recognitions. Once four repetitions of the trained lists had been completed, each of the other lists were read once in order in a test for interlist false recognitions. This procedure was repeated for all nine lists. Errors (both false recognitions and rejections) were recorded.

The results revealed 88.4 percent overall recognition accuracy across all lists; 87 percent on intralist comparisons and 89.9 percent on interlist comparisons. Phrases had 87.0 percent and 95.8 percent accuracy for intralist and interlist comparisons, respectively. Key words had relatively poor recognition – 79 percent and 87.5 percent for interlist and intralist comparisons. The numerals, however, were recognized with 100 percent accuracy. Heading (course) and glidepath advisories (lists 1, 2, 3, and 4) were discriminated from information advisories (list 5) 92.5 percent of the time. Of interlist errors, 23 percent were due to rejections. (A rejection is an event in which the VIP-100, in effect, says 'I don't recognize the spoken phrase.')

While these scores are lower than values required by the SUS, several factors should be considered:

a. The VIP-100 used in the tests was an 'early' version made available for showroom demonstration. A newer version with increased capability is available.

b. The VIP-100 is sensitive to changes in pitch and inflection. This sensitivity is reduced on newer versions by virtue of basic feature changes.

c. The VIP-100 system tested utilized only 16 time samples. This can be increased to 32.

d. The recognition system was 'self-standing.' Additional software which compares expectancies with 'observed' speech inputs, for example, can be added for the training system application.

TABLE 6. VIP-100 TEST PHRASES

List – I	List – II
Approaching glidepath	And holding
Begin descent	And coming down
On glidepath	And coming rapidly down
Going above glidepath	And coming slowly down
Slightly above glidepath	And going further above
Well above glidepath	And correcting
Going below glidepath	And coming up
Slightly below glidepath	And coming rapidly up
Well below glidepath	And coming slowly up
Going rapidly below glidepath	And going further below

TABLE 6. VIP-100 TEST PHRASES (Cont)

List – III	List – IV
On course	Heading is 247
Going right of course	Heading is 245
Going left of course	Heading is 249
Right of course	Heading is 159
Left of course	Heading is 199
Slightly right of course	Heading is 195
Slightly left of course	Heading is 012
Well right of course	Heading is 201
Well left of course	Heading is 102
Turn right, heading	Heading is 360

List – V	List – VI
Wind is 210	Glidepath
Centerline is right	Heading
Cleared to land	Course
Runway 24 right	Well
At decision height	Slightly
One mile to touchdown	Above
Over approach lights	Below
Over landing threshold	On
Execute missed approach	Right
Contact tower	Left

TABLE 6. VIP-100 TEST PHRASES (Cont)

List – VII	List – VIII
Holding	And
Coming	Of
Going	From
Rapidly	Is
Slowly	If
Further	At
Correcting	One-half
Centerline	Up
Approaching	Down
Approach	Half

List – IX
One
Two
Three
Four
Five
Six
Seven
Eight
Nine
Zero

GLOSSARY

AFT	Automated flight training
ARPA	Advanced Research Projects Agency
ASCS	Adaptive Syllabus Control Subsystem
ASR	Automatic speech recognition
ATE	Automated training evaluation
AZ	Azimuth
C	Criteria value
CIC	Combat information center
C_N	Criteria for performance
CRT	Cathode ray tube
D	Difficulty level
EL	Elevation
FAA	Federal Aviation Agency
GCA	Ground controlled approach
GCI	Ground controlled intercept
NAS	Naval Air Station
NATTC	Naval Air Technical Training Center
NAVTRAEQUIPCEN	Naval Training Equipment Center
NM	Nautical mile(s)
NTIS	National Technical Information Service
PAR	Precision approach radar
PTI	Perception Technology, Inc.
R	Run number
S	Score
σ	Standard deviation
SUS	Speech understanding system
TTI	Threshold Technology, Inc.